DOCUMENT RESUME

ED 258 979 TM 850 255

AUTHOR Turner, Susan D.; And Others
TITLE Developmenal Writing Program: Annual Evaluation Report.
INSTITUTION Hillsborough County Board of Public Instruction, Tampa, Fla.
PUB DATE Sep 84
NOTE 128p.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS Academic Achievement; Attitude Measures; Evaluation Methods; Inservice Teacher Education; Intermediate Grades; Pretests Posttests; *Program Effectiveness; Program Evaluation; *Program Implementation; Student Attitudes; Teacher Attitudes; Teaching Methods; Writing (Composition); *Writing Instruction; *Writing Skills
IDENTIFIERS *Developmental Writing Program; *Hillsborough County Public Schools FL

ABSTRACT

The Hillsborough County (Florida) Developmental Writing Program (DWGP) was implemented in grades four to six during the 1983-84 school year. The DWP was evaluated to determine the degree of classroom implementation and program effectiveness. Many instructional innovations resulted from program implementation. Students exhibited improved writing skills as a result of participation in the program as evidenced by pretest and posttest scores. The program's effectiveness was reflected by students' higher achievement test scores in reading comprehension, language mechanics, and language expression. Teachers and students developed more positive attitudes toward writing as a result of their participation in the DWP. This final evaluation report consists of a program description, evaluation design, procedures, analysis and results, and conclusions and recommendations. Appendices include DWP control group end of year survey; teacher interview; program topics and procedures; teacher and student attitude surveys; interim report; holistic score scale; list of raters; writing sample growth charts; frequency distributions; inservice training component; and levels of use chart. The DWP Scoring Manual is attached. (DWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED258979

Developmental Writing Program

Annual Evaluation Report

September, 1984

Prepared by:

Susan D. Turner, Ph.D.
Supervisor of Evaluation
Department of Testing and Evaluation
John A. Hilderbrand, Ph.D., Director
Hillsborough County Public Schools

Kathryn Kaney, Consultant
Planning, Development and Evaluation Associates
Tampa, Florida

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
S. D. Turner

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TM 850 255

EXECUTIVE SUMMARY

During the 1983-84 school year, Hillsborough County's Developmental Writing Program (DWP) was implemented at grades 4 through 6 in thirty-five (35) elementary schools throughout the district. Because of the experimental nature of the program, an extensive evaluation was conducted. The Department of Testing and Evaluation, at the request of the Department of Elementary Education, evaluated the degree of classroom implementation and the effectiveness of the Developmental Writing Program.

Overall. The Developmental Writing Program, which operated in 35 schools in grades 4-6, involved 98 teachers and approximately 3,000 students during the 1983-84 school year. The program appeared to be implemented as intended. Its implementation resulted in improved teacher and student attitudes toward writing and improved student writing. Most teachers changed their instructional behaviors as a result of the Developmental Writing Program.

Implementation. Almost all participating teachers felt they were adequately trained to teach the Developmental Writing Program. Although sufficient instructional time and materials were cited as program weaknesses, most teachers instructed students for at least one and one-half to two hours per week in writing skills using the Developmental Writing Program approach. Almost all teachers were enthusiastic about the Developmental Writing Program, and students reportedly enjoyed and benefitted from this instruction.

Students' Pre to Post Writing. At least 70 percent of students who participated in the Developmental Writing Program exhibited improved writing skills after eight months of Developmental Writing Program instruction. Fifty-seven (57) percent of students who did

not participate in the Developmental Writing Program, but who may have received other types of writing instruction, also exhibited improved writing skills after eight months in school. However, when compared with these other instructional methods, the Developmental Writing Program was significantly more effective in improving students' writing skills.

Students who were instructed by first-year Developmental Writing Program teachers made about the same progress in writing as those instructed by second-year Developmental Writing Program teachers.

Students' Achievement Test Scores. The Developmental Writing Program was at least as effective as more traditional types of language arts instruction in teaching students basic language skills. In some instances, Developmental Writing Program instruction appeared more effective than traditional instructional methods (e.g., language books, worksheets). Fourth, fifth, and sixth grade students who participated in the Developmental Writing Program scored at least as well as a comparable group of non-participants in reading comprehension, language mechanics, and language expression.

Students' Attitudes Toward Writing. Students who participated in the Developmental Writing Program and in other types of writing instruction had more positive attitudes toward writing at the end of the school year than at the beginning of the year. However, Developmental Writing Program participants were significantly more positive in their attitudes toward writing than the comparable group of non-participants.

Teachers' Attitudes Toward Writing. One group of first-year Developmental Writing Program teachers were significantly more positive in their overall attitudes toward writing at the end of the school year than a comparable group of non-participating teachers. All first-year Developmental Writing Program teachers had significantly more positive attitudes toward the teaching of writing than teachers in the non-participating group.

Teachers' Instructional Behaviors. Almost all participating Developmental Writing Program teachers changed their instructional behavior in the area of writing. More than half the total group of first and second-year teachers implemented the program in a "mechanical" fashion. That is, most effort focused on managing instructional time, organizing materials, finding topics, determining methods for editing and publishing, and learning about the requirements of the program. About a third of Developmental Writing Program teachers implemented the program on a "routine" basis. That is, they were comfortable and consistent in their approach to the program. For these routine users, time, materials, topics, editing, publishing and other routines were clearly established; writing classes were planned and operated with a minimum amount of stress. These percentages are as expected during the initial phases of a new instructional program.

TABLE OF CONTENTS

0.0 Introduction..... 1

1.0 Program Description..... 2

2.0 Evaluation Design

2.1 Process Investigation..... 5

2.2 Product Investigation..... 6

3.0 Procedures

3.1 Process Procedures..... 10

3.1 Product Procedures..... 10

3.2.1 Paper Preparation..... 11

3.2.2 Pre-Scoring..... 11

3.2.3 Rater Training..... 16

3.2.4 Scoring..... 19

4.0 Analyses/Results

4.1 Process..... 22

4.2 Product..... 26

5.0 Conclusions/Recommendations

5.1 Conclusions..... 51

5.2 Recommendations..... 54

References..... 55

Attachment 1: Developmental Writing Program Scoring Manual..... 95

APPENDICES

Appendix A: Developmental Writing Program -- Control Group End-of-Year Survey..... 56

Appendix B: Developmental Writing Program Teacher Interview... 58

Appendix C: Developmental Writing Program Topics and Procedures..... 61

Appendix D: Developmental Writing Program Teacher (30-items) and Student (15-items) Attitude Surveys..... 67

Appendix E: Developmental Writing Program Interim Report -- November 1983..... 70

Appendix F: Developmental Writing Program Focused Holistic Score Scale..... 76

Appendix G: Developmental Writing Program List of Raters..... 78

Appendix H: Developmental Writing Program Writing Sample Growth Charts and Pre-Post Differences by Grade... 80

Appendix I: Developmental Writing Program Frequency Distributions..... 84

Appendix J: LoU Chart..... 89

Appendix K: Developmental Writing Program Inservice Training Component..... 92

List of Tables

Table 1: Developmental Writing Program Subjects..... 8

Table 2: Developmental Writing Program Pre-Scored Papers by Agreement Category..... 15

Table 3: Developmental Writing Program Pre-Scored Papers -- Distribution of True Scores..... 16

Table 4: Summary of Developmental Writing Program Student Responses 25

Table 5: Significance of Pre to Post Differences for Each Treatment Group on Writing Samples..... 30

Table 6: Pre to Post Differences Among Treatment Groups on Writing Samples..... 32

Table 7: CTBS Scale Scores (SS) Differences by Grade Level and Treatment Group..... 35

Table 8: Summary of CTBS Scale Score Differences by Grade Level.... 36

Table 9: Developmental Writing Program and Control Group Pre to Post Mean Scores on Student Attitude Survey..... 38

Table 10: Two Developmental Writing Program and Control Groups' Pre to Post Mean Scores on Teacher Attitude Survey -- Total Score and Two Partial Scores..... 41

Table 11: Developmental Writing Program Levels of Use Assignments -- April 1984..... 48

Table 12: Developmental Writing Program Levels of Use Assignments by Teacher Group -- April 1984..... 50

List of Figures

Figure 1: Developmental Writing Program Average Pre and Post Scores for Each Treatment Group..... 28

Figure 2: Developmental Writing Program and Control Group Pre and Post Mean Scores on Student Attitude Survey..... 39

Figure 3: Two Developmental Writing Program and Control Groups' Pre to Post Mean Scores on Teacher Attitude Survey -- Total Score..... 43

Figure 4: Two Developmental Writing Program and Control Groups' Pre to Post Mean Scores on Teacher Attitude Survey -- Attitude Toward Own Writing..... 44

Figure 5: Two Developmental Writing Program and Control Groups' Pre to Post Mean Scores on Teacher Attitude Survey -- Attitude Toward Teaching Writing..... 45

0.0 INTRODUCTION

During the 1983-84 school year, Hillsborough County's Developmental Writing Program (DWP) was implemented at grades 4 through 6 in thirty-five (35) elementary schools throughout the district. Because of the experimental nature of the program, an extensive evaluation was conducted. The Department of Testing and Evaluation, at the request of the Department of Elementary Education, evaluated the degree of classroom implementation and the effectiveness of the Developmental Writing Program.

Preliminary evaluation of the effects of the DWP pilot program, conducted in 1982-83 at 11 elementary schools, indicated that fourth and fifth grade students who participated in one semester of the DWP significantly increased their writing skills. Although results of this one-group pretest-posttest study were promising, program decision-makers wished to draw more rigorous conclusions about the effects of their program. Thus, an equivalent control group (students not involved in the DWP) was included in the evaluation study during the full 1983-84 academic year.

This document serves as the final report for the 1983-84 Developmental Writing Program, and is presented in five sections: Program Description, Evaluation Design, Procedures, Analysis/Results, and Conclusions/Recommendations.

1.0 PROGRAM DESCRIPTION

Early in the 1982-83 school year, instructional supervisors, Ms. Jackie Matthews and Dr. George Mitchell, sought to identify and develop an effective model for an elementary grades writing program. This model was intended to expand the traditional language arts program to include the teaching of writing skills. Eventually, the program, piloted in 11 schools with 21 teachers during the second semester of 1982-83, is intended to be implemented in all district elementary schools.

After researching a variety of approaches to teaching writing, Ms. Matthews determined that a systematic, developmental approach would best meet local needs. The program she designed was based on the belief, currently supported by theorists in the field (Emig, 1971; Graves, 1973; Murray, 1968) that composing is a process rather than a product and that writing instruction should address the entire process including prewriting, drafting, editing and revising, and sharing.

- Pre-writing, the initial stage of the writing process, precedes the act of composing. Pre-writing activities are used to assist students in developing sensitivity to language, to generate student-selected topics and to develop an organizational framework or plan for writing.
- Draft writing, the second stage of the writing process, includes the first and all improved drafts which precede the final or published copy. Draft form, a technique in which students write on every other line, is used in writing each draft copy. This technique enables the writer to transcribe thoughts freely and fluently without concern for form or convention. In addition, the skipped lines facilitate improvement without erasure. Thus, the writer retains his original ideas throughout the draft process. This retention of the writer's ideas expedites transition to the editing and revision process.

- The student then edits by correcting technical errors in spelling, mechanics, sentence structure and usage. Next, by adding, deleting, shifting and combining words, phrases or sentences, the writer achieves stylistic improvements which result in a clear, cohesive and unified composition.
- Sharing, or "publishing", completes the process and occurs when the student communicates the writing product to others by reading it aloud, posting it on a bulletin board, or contributing it to a formally published collection.

In order that the program be manageable by a large number of teachers working with students at different grade and ability levels, Matthews and Mitchell decided that it must be a total integrated curriculum rather than merely a collection of writing activities. Therefore, writing instruction included four dimensions: words, sentences, paragraphs and longer discourse, at four levels (not necessarily grade levels) of developmental competency. (See The Developmental Writing Program Teacher Guide Grades 3-6 for detailed explanations of each dimension and level.)

Thus far, the theoretical basis for the program, its instructional techniques and management strategies have been presented to approximately 110 participating teachers at 35 elementary schools in the district. Training has occurred for three separate groups of teachers:

1) In 1982-83, 21 pilot teachers in 11 elementary schools received three days of inservice spaced between January and April 1983. These teachers are heretofore referred to as the "Stage I" group.
2) In 1983-84, 42 additional teachers in the original schools received three days of inservice spaced between August 1983 and February 1984. These teachers are heretofore referred to as the "Stage IA" group.
3) Also in 1983-84, 45 teachers in 24 additional schools received three consecutive days of inservice in August 1983. These teachers are heretofore referred to as the "Stage II" group.

Given broad guidelines as well as specific training which they could tailor to their individual situations, teachers were expected to apply their training in regularly occurring (at least three times weekly) writing lessons in their classes. Program supervisors conducted the training sessions and visited classrooms to observe and assist teachers. Teachers were provided with instructional filmstrip/tape kits on teaching writing, a manual by Dr. Evelyn Rothstein entitled Teaching Writing: A developmental Systematic Approach and one copy of the appropriate levels of its accompanying workbooks, poster kits, and other reference books.

2.0 EVALUATION DESIGN

The evaluation approach followed a process-product model. Process information was collected via periodic interviews and surveys. The product portion of the model was accomplished using a control-group pretest-posttest design. This section will delineate evaluation questions and subordinate investigations that were addressed by the evaluation and will describe the subjects and instruments that were used.

2.1 Process Investigation

A descriptive summary of the degree of implementation of the DWP, including descriptions of classroom activities, training sessions, and support services were provided to program decision-makers. Process questions included such issues as:

- What types of training activities were offered for the DWP? Were teachers and other implementors trained? Was the training adequate? Are training sessions in need of modifications?
- Is the DWP being implemented as intended? How often does instruction take place? What are student reactions to the program?
- Have teachers modified the program delivery in any way?
- Are there sufficient materials for implementation of the program?
- Which specific features of the DWP are most used and/or most effective? Which features are least used and/or least effective?
- Do DWP teachers at the same site collaborate in planning for and/or instruction in writing?
- How does the DWP impact the rest of the language arts curriculum?
- What kind of writing instruction do control student subjects receive?

To respond to these questions, all DWP teachers an a control group of teachers contributed process information. Teachers who participated in DWP training sessions were surveyed following each session. In addition, a stratified sample of DWP teachers were observed and/or interviewed during the academic year. Control group teachers completed an end-of-the-year survey. (See Appendix A.)

Interview, observation, and survey forms addressing the process questions were developed by the Departments of Testing and Evaluation and Elementary Education. In addition, records were kept of training sessions and other types of DWP activities. A modified form of the Levels of Use (LoU) system (Hall, Loucks, Rutherford, and Newlove, 1975) was used as the interview schedule. (See Appendix B.)

2.2 Product Investigation

The following product questions were addressed by the study:

1) Is student writing better after DWP instruction?

2) Do students who receive instruction in the DWP write better than students who receive a different type of writing instruction?

3) Do students who receive instruction in the DWP perform as well as other students on tests measuring traditional language conventions?

4) Is student attitude toward writing different after DWP instruction?

5) Is teacher attitude toward writing different after DWP participation?

6) Are teachers' instructional behaviors changed as a result of training received in the DWP?

Questions 1 and 2, the primary product issues, are discussed first.

Because the DWP is delivered under three distinct sets of conditions, the district wished to investigate whether different results were associated with different delivery conditions. In addition to primary product questions 1 and 2, therefore, the following additional question also was investigated:

. Among students enrolled in classes where the DWP is implemented, is there differential performance across the following groups:

> Experimental Group Stage I (E1) - Students enrolled in classes with teachers who are in their second year of the program.
>
> Experimental Group Stage IA (E2) - Students enrolled in classes with teachers in their first year of the program, who are being assisted by second-year teachers.
>
> Experimental Group Stage II (E3) - Students enrolled in classes with teachers in their first year of the program, not assisted by second-year teachers.

To address primary product questions 1 and 2 and their subordinate investigation, writing samples were collected from students who represented each of the groups defined by the experimental groups described above. In addition, writing samples also were collected from a Control Group (C), students enrolled in classes where the program was not being implemented. A sampling plan was developed to ensure representation of twelve groups of students, as illustrated in Table 1. Each cell was represented in proportion to the number of students who actually received DWP instruction at each grade level. In all, 775 subjects constituted the sample: 180 subjects in Stage I, 198 subjects in Stage IA, 198 subjects in Stage II, 199 subjects in the Control Group.

Table 1

Developmental Writing Program Subjects

Grade 4	Grade 5	Grade 6
E1	E1	E1
E2	E2	E2
E3	E3	E3
CG	CG	CG

Two types of instruments were used: test instruments and evaluation instruments. Each type is described in this sub-section.

Tests. Two samples of student writing were collected from all subjects during the first two weeks of the 1983-84 school year. In May 1984, two more writing samples were collected from all subjects. Altogether, two topics were administered. On each occasion, however, students were assigned a single topic. On the pretest occasions, therefore, students wrote on each of the two topics, and again on the posttest, each wrote on the same two topics. In all, each student wrote four papers.

Topics were developed by Hillsborough County instructional personnel, and administration procedures were developed by Hillsborough County Testing and Evaluation personnel. A copy of each topic and procedures can be found in Appendix C.

Evaluation. A scoring manual was developed to train raters in a form of general impression scoring called focused holistic scoring. This manual was used to guide the evaluation of all papers. A copy of the Developmental Writing Program Scoring Manual accompanies this report as Attachment 1.

Question 3 was addressed by utilization of the district's achievement test (CTBS). Scale scores in the areas of reading comprehension, language mechanics and language expression from the spring 1984 administration of the CTBS were obtained for each subject in the sample.

Questions 4 and 5 were addressed via completion of attitude surveys. Teacher and student attitude surveys, developed and fieldtested by the Department of Testing and Evaluation during the first evaluation stage (1982-83), were used to measure attitudes toward writing and the teaching of writing. (See Appendix D for a copy of each survey.)

All teachers in Stages IA and II and the Control Group were administered a teacher attitude survey prior to the first DWP training session. All participating and control students were administered a student attitude survey during the first two weeks of the 1983-84 school year prior to any instruction in the DWP or other writing program. For post-information, in May 1984, the same teachers again completed the same teacher attitude survey; students were administered the same student attitude survey.

Question 6 was addressed by using an adaptation of the Concerns Based Adoption Model Levels of Use interview system. A stratified sample of DWP teachers were interviewed in March 1984 to determine whether their instructional behavior had changed as a result of the DWP.

3.0 PROCEDURES

3.1 Process

Using a session-specific instrument, each DWP training session was evaluated at the time of its occurrence. Following analysis, verbal information was shared with program supervisors for the purpose of improving future training offerings.

In November 1983, independent contractors under the direction of the Department of Testing and Evaluation conducted interviews and observations with a sample of teachers in Stages I, IA and II. Verbal and written reports of findings were shared with program supervisors in December 1983. (See Appendix E.)

3.2 Product

Each writing sample was assigned a score from "1" to "4", with "1" constituting the low end of the scale. Pre to post-writing sample scores were compared for all groups and between groups using a Wilcoxon Matched-Pairs Signed-Ranks Test. Data first were compared for all subjects in each group, then factored according to grade level. Significance levels were set at p .05 for all analyses.

In January 1984, the consultant, Ms. Kaney, met with Drs. Hilderbrand and Turner to discuss evaluation activities related to scoring the writing samples. Another planning meeting was held in February. Subsequent project activities were conducted in five sequential stages: Paper Preparation, Pre-Scoring, Rater Training, Scoring, and Follow-Up. Each stage will be discussed in detail.

3.2.1 Paper Preparation

Following the February planning meeting, the consultant prepared a logistics document to guide all paper preparation and other scoring-related activities. The logistics document provided instructions for the following activities:

1) Randomly selecting papers from each of twelve groups to be included in the scoring sample.
2) Recording a seven-digit numerical code on each paper, then removing all identifying information from the paper.
3) Scrambling paper order to interrupt all unit clusters.
4) Pre-coding a scannable scoring form for each paper.
5) Preparing scoring packets that contained a designated number of papers and their corresponding scoring forms.
6) Preparing ancillary materials for the scoring session.

The logistics document also detailed procedures and staff responsibilities for training, scoring, and follow-up activities.

3.2.2 Pre-Scoring

It can be claimed, with reason, that anyone who teaches writing should know how to "grade" or score the papers that students write. Not only should they be able to do so, most are able to do so. Nonetheless, teachers often are startled, in discussions with their equally well-trained and conscientious peers, to learn that their colleagues do not always share their composition priorities. Some teachers place greater emphasis on neatness than others. Some are more concerned with the structural correctness of sentences; others are more concerned with the richness of language in sentences. Some teachers place greater emphasis on a composition's content; others place greater emphasis on a composition's form. Although all of these priorities are worthy ones, they can be the root of considerable disagreement among raters as to the score that a given paper should receive, thereby resulting in unreliable and invalid scores.

To counteract the influences of these conflicting priorities during scoring activities, a single standard, established by expert judges who represent the district's priorities, is established in advance through pre-scoring activities. Pre-scoring is done to establish the standard for all ensuing scoring and to ensure the validity both of the process and of the scores ultimately given to each paper. During pre-scoring, a team of expert judges who represent the system's present standards undergo intensive training. They then score, discuss, and rescore until they have reached agreement about the final score on a large number of papers. These pre-scored papers then are used to stabilize the score scale throughout subsequent training and scoring activities.

Some pre-scored papers become Anchor Papers, so called because they anchor or define the score scale for training purposes. Raters study these papers, along with the score scale, early in the training process. Another set of pre scored papers become Training Papers and are used to teach raters to score in concert with the pre-scoring standard, rather than apply personal standards of their own. Still another set of pre-scored papers become Check Papers. These papers are used periodically throughout scoring as a check to ensure that raters are not drifting away from the standard set during pre-scoring.

To conduct pre-scoring for this program, 100 papers were representatively selected from those submitted across all 12 groups. These papers were blinded and pre-sorted such that all unit clusters were interrupted. The papers then were organized for pre-scoring.

Pre-scoring activities were conducted on June 13 in the district Testing and Evaluation office. Three pre-scoring judges were selected as follows: one school-based curriculum specialist, one instructional supervisor, one measurement/evaluation specialist. The consultant designed and directed pre-scoring activities.

The consultant gave each judge a copy of the Scoring Manual and reviewed the manual's introduction and the sections related to hand-scoring issues and procedures and regulations.

Each of the two topics and their administration instructions also were reviewed. The consultant then introduced the judges to focused holistic scoring by carefully reviewing the section, "Expectations for the Composition," which defined each of the four composition domains and identified each of the specific traits to be examined. The consultant explained to the judges that each paper would be assigned only one score by each judge, but that this score was to reflect the student's general performance across all four domains on the specified traits.

Once this section had been reviewed, judges studied the score scale that they were to use. (A copy of "Developmental Writing Program-Focused Holistic Score Scale" is included as Appendix F.) To begin pre-scoring activities, each judge was given a set of copies of the same five papers. Judges were asked to read each paper, compare the paper's quality with each of the descriptors on the score scale, then select the descriptor most appropriate for the paper, compare the paper's quality with each of the descriptors on the score scale, then select the descriptor most appropriate for the paper and assign the associated score to the paper. Judges were instructed to read and score each of the five papers independently, without discussion with other judges and without regard for any other paper in the set.

Once each judge had read and scored all five papers, the consultant collected all scores. First, papers on which all judges had assigned the same score (firm-agreement papers) were identified, one at a time. For each of these papers, the consultant gave the paper's identification number and the judges' score, now designated the "True Score." Judges then were asked to re-read the paper and the descriptor to firmly establish the rationale for the paper's true score. Judges briefly summarized the true score's rationale on the back of the score card.

Once this procedure had been followed for each of the firm-agreement papers, judges again reviewed these papers, their scores and rationale statements. Judges then were asked to re-read the remaining papers and rescore them, utilizing

the clarification that had resulted from their review of firm agreement papers. After this second scoring, results were collected and the above procedures were repeated for all papers on which there had been firm agreement.

Papers which still had discrepant scores were then identified and the following procedures were employed to resolve the disagreement.

1) The consultant identified the scores that judges had assigned to the discrepant paper. Usually, only two scores were represented in the disagreement. For example, two judges may have assigned a score of "3" and one judge may have assigned a score of "4." In 2 percent of the cases (two of the 100 papers) each judge had assigned a unique score to a single paper.

2) The consultant then referred judges to firm-agreement papers which represented each of the assigned scores and judges were asked to compare the discrepant paper to these papers. Judges were to determine which set of papers the discrepant paper was most like. This comparison was followed by a discussion of the discrepant paper's similarity to other papers. After the discussion, the consultant called for a third scoring of the discrepant paper. This third scoring resulted in a firm score in only .03 additional cases. This small increase in agreement after the third reading was in concert with the consultant's experience in conducting other iterative decision-making activities. Little-to-no movement is expected in judges' decisions beyond this point. The consultant, therefore, resolved the final scores for the sixteen remaining papers by defining the true score as the score given by two of the three judges.

To reflect the differences in ease in establishing true scores, pre-scored papers are grouped into four categories. The categories and their definitions follow:

A) Firm-Agreement Papers - pre-scored papers on which all judges assign the same score after the first reading.

B) Second-Round-Agreement Papers - pre-scored papers on which judges assign the same score after the second reading.

C) Third-Round Agreement Papers - pre-scored papers on which all judges assign the same score after the third reading.

D) Resolved-Agreement Papers - pre-scored papers which remain discrepant after three readings and which are resolved by a designated referee.

All four types are useful in training raters because they help to ensure that raters encounter not only papers that are easy to score (categories A and B), but papers that are difficult to score (categories C and D) as well. By encountering difficult papers during training, raters, like judges, are able to fine-tune their conceptions of the score scale.

Table 2, below, displays the proportion of pre-scored papers that were grouped into each of the four categories as a result of pre-scoring.

Table 2

Developmental Writing Program Pre-Scored Papers by Agreement Category

(N = 100)

Category A	Category B	Category C	Category D
.61	.20	.03	.16

The true score distribution among pre-scored papers is shown in Table 3.

Table 3

Developmental Writing Program Pre-Scored Papers
Distribution of True Scores

(N = 100)

Score 1	Score 2	Score 3	Score 4
.08	.41	.45	.06

3.2.3 Rater Training

Once pre-scoring had been completed, the consultant selected Anchor Papers, Training Papers, and Check Papers, in preparation for rater training. For each type, the following considerations were made:

Anchor Papers: It is imperative that the same number of anchor papers be selected to represent each scale value. This consideration prevents raters from establishing any expectation regarding score distribution. Because few scores of 4 and 1 were represented in the pre-scored sample, the number of anchor papers representing each scale value was set at three. For each scale value, the three anchor papers were selected such that one of them represented firm-agreement papers, one represented second-round-agreement, and one represented either third-round-agreement or resolved agreement. This combination was designed to illustrate the range of performance that judges found in each scale value.

Training Papers: To the extent possible, training paper sets also represented a full distribution of scores. Each of the five training sets contained five papers, selected such that at least three of the four scale values were represented. In each set, however, there were more papers representing the central score values (3 and 2) than the outlying values (4 and 1). The over-representation of central values in each set was designed to give raters extensive training in differentiating between these scores, traditionally the most difficult values for raters to separate. Again, as with anchor paper selection, training paper sets also were selected such that all possible agreement categories were represented in each set.

Check Papers: These sets, each containing five papers, were selected using the same considerations as those used to select training papers.

Once each of the above paper types had been selected, sufficient copies of each set were made for each rater, and sets were collated and labeled for use in rater training and scoring activities.

Training and scoring activities were conducted in the Town and Country Elementary School library, on June 26, 27, and 28. Twelve raters were selected by district staff and were representative of schools throughout the district. Both DWP teachers and non-DWP teachers were represented among raters. A list of raters and their school affiliations is shown in Appendix G.

Training activities required approximately six hours. Raters were first given a review of the DWP and were given brief information about the make-up of the scoring sample. They were told that they would be scoring papers both of students in the program and of students not in the program. They also were told what grade levels were represented in the scoring sample and that both pretest and posttest papers were represented. The consultant explained to raters that all identifying information had been removed from the papers and that the papers had been scrambled so that there would be no grade-level, program, group, or occasion clusters in their packets.

Each rater then was given a copy of the Scoring Manual and each section of the manual was carefully reviewed with them by the consultant. Raters were given time to review the "Expectations" section, the score scale and the anchor papers, and to discuss these materials among themselves and with the consultant until they felt reasonably comfortable with them.

Raters then were given papers for Training Round #1 (T-1). Each rater was given a set of the same five papers and a score sheet on which to record the papers' scores. Raters red and scored each paper independently. They were instructed to utilize the score scale in combination with the

anchor papers to determine the score for each of the papers in T-1.

Once raters had finished scoring papers in T-1, their score sheets were collected and their scores for each paper recorded on the consultant's monitor sheet. Once all raters' scores had been transferred to the monitor sheet, their scores were compared to each paper's true score, and calculations were made to determine the number of discrepant scores for each rater (incidence) and whether their discrepancies tended to be high or low as compared with the true scores (tendency). In cases where raters' incidence or tendency indicated a problem, this information was related to them and recommendations were made to them for adjustment of their scoring. Otherwise, their scoring was simply monitored to assess individual accuracy. This score monitoring continued throughout all training rounds and check rounds.

Once each raters' accuracy had been evaluated, raters' score sheets were returned to them along with a self-adhesive label that displayed the true score for each paper in the set. Raters were instructed to affix the label to the front cover of the paper set and to review these true scores. Where their score was discrepant from the true score, raters were instructed to utilize anchor papers and other training papers, along with discussion to clarify the discrepancy. In cases where as many as one-third of the raters were discrepant on a single paper, the consultant discussed the paper with the entire group. Raters were not allowed to dispute or otherwise challenge the true score.

When raters had completed their discussions, they were given another set of training papers and the procedures described above were repeated. Altogether, five training rounds were conducted.

Monitor sheets for each training round document quite strong agreement among raters throughout training activities. Both tendency and incidence ratings were assigned to each rater. Positive and negative tendency ratings indicate whether a rater was scoring higher or lower than the true score, respectively. The value of the tendency rating indicates how

many points above or below the true score the rater scored for the entire round.

The incidence value represents the number of discrepant scores assigned by the rater for the round. An incidence rating of zero indicates that the rater assigned no discrepant scores in that round. A rater who had an incidence rating of 2 or higher, but had a tendency rating of 0, showed discrepancy from the true scores, but no consistent tendency to overscore or underscore. Of the 300 scores assigned throughout the five training rounds, only four cases (.01) of non-adjacent discrepant scoring were observed. That is, discrepant papers were more than one point away from the true score in only four cases.

3.2.4 Scoring

Training activities were completed on the first day of the session. Because there had been an overnight break between the final training round and the first scoring round, a check round was conducted before the first scoring round began. This and all ensuing check rounds were conducted in the same manner as the training rounds had been.

Each scoring round was conducted in exactly the same manner. Each rater was given a scoring packet, containing a set of papers and their corresponding scoring forms. Raters were instructed to first check the order of their materials to ensure that they were recording information on the correct form for each respective paper. Raters then read each paper independently and coded in the score and their rater identification number on the appropriate score form. When raters had finished scoring the entire packet of papers, the packet was collected by a scoring assistant and the rater was given another packet of papers to score. Collected packets were checked to ensure that each paper had been properly coded in. When each score form in a packet had been checked for accuracy and completeness, the score on each form was covered with a WRITE GUARD rating shield to obscure it from view by the second rater. All papers and score forms then were placed in identical order and returned to their packets.

When scoring assistants collected the second packet from raters, they gave raters a packet that already had received a first reading. Raters followed their routine rating procedures, this time recording their identification numbers and scores on the section of the form designated for the second reading.

When all distributed packets had received two readings, a new scoring round began. This procedure continued until all papers had been scored by two readers.

When packets were collected from the second reading, each packet was checked for completeness and accuracy. The rating shield was then removed from each score form and the two scores were compared. When the two scores were found to be discrepant by more than one value, both scores were covered by rating shields. Such score forms were placed on top of the others and these packets were set aside for resolution scoring.

After all scoring packets had received two readings, packets which contained papers with discrepant scores were distributed for resolution. Resolution readings were conducted exactly as other readings had been, the rater recording score information in the area of the score form designated for the resolution scoring.

When all scoring activities had been completed, rating shields were removed from score forms that had required resolution, and each form was check carefully for accuracy and completeness. All score forms were then cleaned of all stray marks, stacked, and returned to the district office for scanning.

Scoring of the writing samples was conducted by Ms. Kathryn Kaney, Consultant, Planning, Development and Evaluation Associates, who developed a four-point focused holistic scoring scale and conducted pre-scoring and scoring sessions.

Average raw scores on teacher and student attitude surveys were analyzed using a students' t-test statistical procedure. CTBS scores were compared using a one-way analysis of variance statistical procedure. (Spring 1983 scores were used to assure initial equivalence among groups.) The same subjects who were selected for the writing sample scoring

(section 3.2.1) also contributed data for the attitude and achievement test questions.

4.0 ANALYSES/RESULTS

4.1 Process

To determine the degree of implementation of the DWP in the classroom setting and to provide timely feedback to program developers about training sessions and DWP occurrences, surveys, interviews and observations were conducted throughout the year with various groups of teachers. After each training session in which teachers completed an evaluation survey, data were tallied by the Department of Testing and Evaluation and results were shared, either verbally and/or in writing, with program presentors.

In general, training sessions were judged at least adequate, and usually above average, in content and delivery. Most teachers reported they gained knowledge and skills from the training sessions. They reported they felt more able to instruct students in writing after attending the training sessions. Some Stage IA teachers, however, stated they were confused by the first workshop day and did not have a solid idea of the program and its purpose until the second workshop day in early October. They suggested restructuring the introductory session to emphasize actual classroom management and materials. (NOTE: Program developers have indeed reorganized the sessions for 1984-85.)

Training request, expressed during interviews, included:

. evaluating/grading students' writing,
. using computers to teach writing,
. expanding one's own writing skills,

. finding places and methods for formal publication of student writing.

During structured interviews, most teachers reportedly provided DWP instruction three or four days per week, a minimum of 30 minutes instruction per day. All teachers who used the DWP stated they regularly used the pre-writing, drafting, editing, and publishing program components. In about a third of interviewed teachers' classrooms, students wrote in their own journals on a frequent basis. Only two teachers (of 30) stated they regularly wrote in a teacher journal.

When asked about particular program strengths they had noted, teachers frequently mentioned the following.

. "The DWP provides a foundation for vocabulary acquisition and development."

. "Children learn to appreciate and be more conscious of words."

. "The DWP provides an avenue for students to apply language skills (punctuation, capitalization, nouns, verbs, etc.)."

. "The program is motivating for students and provides a valuable means of expression for them."

. "DWP skills transfer to other classes, such as science and social studies."

. "The DWP encourages teachers to use creative and unconventional methods of instruction, if they so choose."

. "Reading children's writing allows the teacher to be more familiar with each child's individual personality."

Stated program weaknesses generally related to time and materials. Some teachers felt the pressure of adding one more "subject" to an already overloaded schedule. Others worried that spending time teaching writing would negatively affect district achievement test (CTBS) scores. Others had 35-40 students per class and could not

personally make contact with every student during writing class. Several teachers requested more program materials, specifically ideas for pre-writing activities and writing topics.

Several sixth grade teachers in the DWP expressed concern that their students should not continue to receive instruction in the writing process when they moved to seventh grade. They suggested a continuous program for grades 4-12.

Of specific interest to program developers was the interaction between Stage I and Stage IA DWP teachers. Because these teachers were housed at the same school, it was anticipated that first year (Stage IA) teachers would request and receive assistance from second year (Stage I) teachers. Although some evidence of this collaboration was found, it was not as formal nor as frequent as anticipated. Indeed, in some locations, no interaction was noted.

Interviews with 30 students about the DWP corroborated much of what teachers had reported. As shown in Table 4, about half of the DWP students stated they wrote or had a writing lesson every day. Other students reported writing every other day or two to three days per week. All interviewed students said they participated in the prewriting and drafting aspects of the program. Almost all students commented they were involved with peer editing, using word books, dictionaries and/or thesauri during the writing process, and publishing final copies of their writing. Less than half of the interviewed students reported being aware of or using a writing center in their classrooms.

Overall, it appeared to interviewers and observers that teachers were indeed implementing the program according to district directives. They and their students seemed to be knowledgeable about the writing process and to have had actual experiences with it.

Table 4

Summary of Developmental Writing Program Student Responses

Type of Question	Yes	No
Do you get to write or have a writing lesson every day?	14	16
When you have a writing lesson do you first talk about what you're going to write?	30	0
Do you write drafts of things?	30	0
Do kids in this class work together in editing each others' writing?	29	1
To help in choosing words for writing, have you started using the thesaurus?	24	6
dictionary?	30	0
other word books?	23	6
Do you feel like you can use them easily?	29	0
Can you use them whenever you want?	30	0
Have you found some new words to use when you write?	27	2
Do you keep your writing in a personal folder?	27	2
Do you use your folder during writing class?	27	3
Is there a special place in your room that's called or used as a writing center?	12	19
Do you ever get to go there and do some writing?	6	6
Do other kids get to go to the writing center?	7	5
Have you had a chance to show off or put up anything you've written yet?	28	2
Do you read what you write to other people?	26	4
Has (principal) been in to see or hear anything you or other kids have written?	18	12

4.2 Product

Six questions constituted the product portion of the evaluation. To respond to these questions, several procedures took place.

- All students completed four writing samples, two pre and two post. A four-point rating scale was constructed; selected pre and post papers were scored by a team of trained raters.
- All students completed two attitude surveys, one pre and one post. After all surveys were scored, a sample of pre and post scores were compared.
- CTBS scores in reading comprehension, language mechanics and language expression were researched and compared for selected students.
- Teachers completed two attitude surveys, one pre and one post. Surveys were scored; then pre and between group comparisons were made.
- Interviews were conducted with a group of teachers to determine changes in their instructional behaviors.

Each question and its results will be presented in this subsection.

To answer the evaluation question:

Is student writing better after DWP instruction?

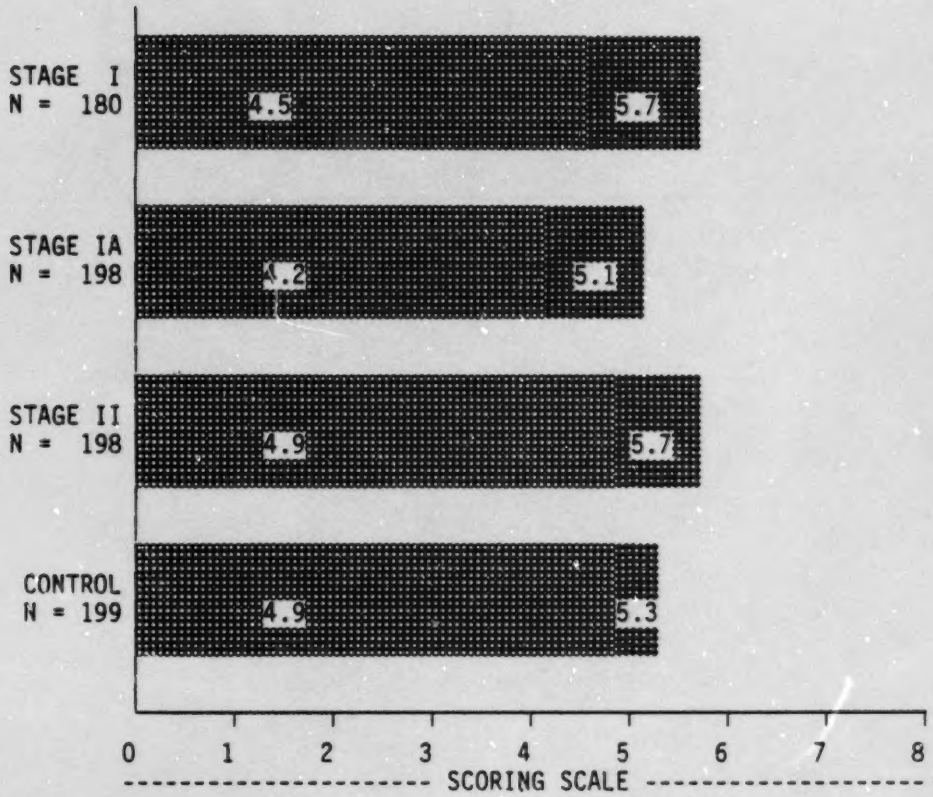
three analyses were conducted. First, each student's scores on the two pre writing samples were summed; then the two post writing samples scores were summed. This allowed average pre and post writing sample scores to be computed for each group. Second, within each group, the percentages of students who earned a lower postscore (negative gain), the same pre- and postscores (no gain) and a higher postscore (positive gain) were computed. Finally, in order to determine each group's growth from pre to post writing assessment, frequency distributions which showed the sum scores from both writing

samples were generated. The following descriptions explain results of the three procedures.

Figure 1 displays average pre and post writing sample scores for each instructional group, the growth experienced by each group during the August 1983 - May 1984 instructional periods, and each group's relative position on the scoring scale. Although students in all groups experienced growth, the three treatment groups outdistanced the control group. Students in Stage I scored an average of 4.5 on the pre writing samples, slightly above the 8-point scale's midpoint. This group's average score on the post writing samples was 5.7, a gain of 1.2 scale points. Students in the Stage IA sample scored an average of 4.2 on the pre samples and 5.1 on the post samples, a gain of .9 scale points. Although both Stage II and Control students scored highest on the pre writing samples, 4.9, they did not experience equitable growth during the year. Stage II students gained .8 scale points for a post writing sample score of 5.7. Control students gained the least number of scale points, .4, for a post writing sample score of 5.3. (Graphs depicting each grade level's pre and post scores are shown in Appendix H.)

Figure 1

Developmental Writing Program Average Pre and Post Scores for Each Treatment Group



AVERAGE SCORE OF PRE WRITING SAMPLES, AUGUST 1983

AVERAGE SCORE ON POST WRITING SAMPLES, MAY 1984
(GROWTH IS INDICATED BY LENGTH OF THIS BAR)

In order to determine whether these within-group pre to post differences were significant a non-parametric statistical procedure, the Wilcoxon Matched-Pairs Ranked-Signs Test, was used. This statistical test is designed for use with ordinal (ranked) data such as the four-point scale used to rate each writing sample. Results of the Wilcoxon comparisons shown in Table 5 indicate that, within each group, students made significant gains in writing skill from the beginning to the end of the school year. Of the 180 Stage I students, 18 scored lower on the post then on the pre writing samples; that is, their writing skills decreased during the school year. Twenty-three students' writing skills remained the same; 139 students (an impressive 77 percent) increased their writing skill during the school year. Results for Stages IA and II resemble those of Stage I with 69 percent and 72 percent, respectively, of students in these groups showing a gain from pre to post assessment.

Significantly, 57 percent of students in the Control group experienced gain in writing skills during the school year. It is important to note that the largest percentage of negative gain, or loss of skill, and the largest percentage of no gain in skills occurred in the Control group. Also noteworthy is the Stage I student performance; the smallest loss of skills percentage and the largest gain percentage occurred in this group which was instructed by second-year DWP teacher.

Pre to post score frequency distributions for the four groups are shown in Appendix I.

Table 5

Developmental Writing Program

Significance of Pre to Post Differences
for Each Treatment Group on Writing Samples

Group	N	Number (%) of students who experienced: Loss (-) from pre to post	No gain (0) from pre to post	Gain (+) from pre to post	Level of Significance
Stage I	180	18 (10%)	23 (13%)	139 (77%)	.00*
Stage IA	198	27 (13%)	36 (18%)	135 (69%)	.00*
Stage II	198	35 (18%)	20 (10%)	143 (72%)	.00*
Control	199	44 (22%)	41 (21%)	114 (57%)	.00*

* denotes statistical significance

30

To answer the evaluation question:

> **Do students who receive instruction in the DWP write better than students who receive a different type of writing instruction?**

an additional statistical procedure, the non-parametric Median Test, was employed. For this study, the median gain score (+1) was computed across all groups; then the percentage of students in each group who exceeded or did not exceed the median was computed. Table 6 shows that comparable percentages of students in the three treatment groups (44, 37, 36 percent) exceeded the median gain score, but significantly fewer students in the Control group (16 percent) exceeded the median gain score. Thus, it appears as if the DWP was more effective in improving students' writing skills than other types of writing instruction delivered in Hillsborough County in 1983-84.

Table 6

Developmental Writing Program

Pre to Post Differences Among Treatment Groups on Writing Samples

		Number of students whose pre to post scores were:	
Group	N	less than or equal to the average difference*	greater than the average difference*
Stage I	180	101 (56%)	79[a] (44%)
Stage IA	198	125 (63%)	73[a] (37%)
Stage II	198	126 (64%)	72[a] (36%)
Control	199	166 (83%)	33[b] (16%)

* Average pre to post difference across all groups = +1 point

[a] These groups are not significantly different from each other.

[b] This group is significantly different from the others.

To answer the evaluation question:

> Do students who receive instruction in the DWP perform as well as other students on tests measuring traditional language conventions?

grade-level (fourth, fifth, and sixth) comparisons of basic language skills were made across groups.

Every spring, all district K-9 students take the CTBS, a standardized achievement test. For a sample of students in each treatment group, 1984 test results were researched in three areas: reading comprehension, language mechanics, and language expression. Average scale scores in each of these subtests were computed for each treatment group. Then, using an analysis of covariance statistical procedure, groups were compared. (In order to account for the possibility of initial non-equivalence among groups, 1983 total reading and total language scores were used as the covariate.)

It was hypothesized there would be no significant difference among groups on any of the subtests. That is, there would be no differential effects for students who had received DWP instruction and those who had received more traditional types of language arts instruction (e.g., language textbook). In particular, program developers, administrators, and teachers were concerned that DWP students would score as well as other students in these areas to show they had indeed received adequate instruction in basic language skills.

Tables 7 and 8 show results of the grade level analyses of DWP and Control group students' CTBS scale scores in reading comprehension, language mechanics, and language expression.

- In grade 4, Stage I DWP students performed significantly better in reading comprehension and language mechanics than students in the Stage IA and Control groups.

- In grade 4, Stage II DWP students performed significantly better in reading comprehension than students in the Stage IA group.

- In grade 4, no significant differences were noted among groups in language expression.

- In grade 5, no significant differences were noted among groups in reading comprehension.

- In grade 5, Stage IA students performed significantly better in language mechanics than students in the Control group.

- In grade 5, Stage II students performed significantly better in language expression than students in the Stage I group.

- In grade 6, no significant differences were noted among groups in any area.

In sum, it appears as if DWP instruction was at least as effective as more traditional types of language arts instruction in teaching students basic language skills. Indeed, in some instances DWP instruction appears more effective than the traditional instructional methods.

Table 7

CTBS Scale Score (SS) Differences by Grade Level and Treatment Group

G R A D E 4

	Stage I (N=76)		Stage IA (N=67)		Stage II (N=74)		Control (N=85)	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Total Reading SS	633.4		599.3		643.3		633.7	
Reading Comprehension		719.6[x]		654.5[#]		718.5[#]		704.5[x]
Total Language SS	648.8		615.7		656.6		645.6	
Language Mechanics		702.0[°]		667.4[°]		696.1		684.9[°]
Language Expression		706.0		664.8		705.6		690.6

[x] These two groups are significantly different.
[#] These two groups are significantly different.
[°] These two groups are significantly different.

G R A D E 5

	Stage I (N=76)		Stage IA (N=67)		Stage II (N=74)		Control (N=85)	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Total Reading SS	684.1		671.9		703.8		695.6	
Reading Comprehension		724.3		724.1		751.1		743.1
Total Language SS	681.7		674.2		695.5		683.9	
Language Mechanics		706.2		707.3[x]		714.6		701.8[x]
Language Expression		696.0[#]		695.0		716.1[#]		705.1

[x] These two groups are significantly different.
[#] These two groups are significantly different.

G R A D E 6

	Stage I (N=76)		Stage IA (N=67)		Stage II (N=74)		Control (N=85)	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Total Reading SS	758.4		704.8		748.1		733.5	
Reading Comprehension		805.0		753.8		793.2		782.2
Total Language SS	726.7		683.8		730.2		713.3	
Language Mechanics		739.8		697.5		739.8		716.9
Language Expression		742.0		701.7		737.9		725.8

Table 8

Summary of CTBS Scale Score differences by Grade Level

	Subtest		
	Reading Comprehension	Language Mechanics	Language Expression
Grade 4	.Stage I scored higher than IA and Control .Stage II scored higher than IA.	.Stage I scored higher than IA and Control	.No difference among groups
Grade 5	.No difference among groups	.Stage IA scored higher than Control.	.Stage II scored higher Stage I.
Grade 6	.No difference among groups	.No difference among groups	.No difference among groups

To answer the evaluation question:

> **Is student attitude toward writing different after DWP instruction?**

two analyses were conducted. First, within-group average pre and post attitude survey scores were computed and compared for a sample of students who were enrolled in any of the DWP stages and a sample of students who were not enrolled in the DWP. Second, average postscores were compared between these two groups. Table 9 shows results of both procedures. Based on results of a t-test for independent samples, the two groups of students were not significantly different at the beginning of the 1983-84 school year in their attitudes toward writing. At the end of the year, however, the DWP group exhibited significantly ($p \leq .000$) more positive attitudes toward writing than the Control Group.

Both groups experienced more positive attitudes toward writing at the end than at the beginning of the year. A score close to 15 indicates a positive attitude toward writing; a score close to 45 indicates a negative attitude toward writing. As shown in Figure 2, both groups of students moved in a positive direction in their attitude toward writing during the school year. However, the steeper growth slope of the DWP group indicates a more substantial positive change.

Table 9

Developmental Writing Program and Control Group
Pre to Post Mean Scores on Student Attitude Survey

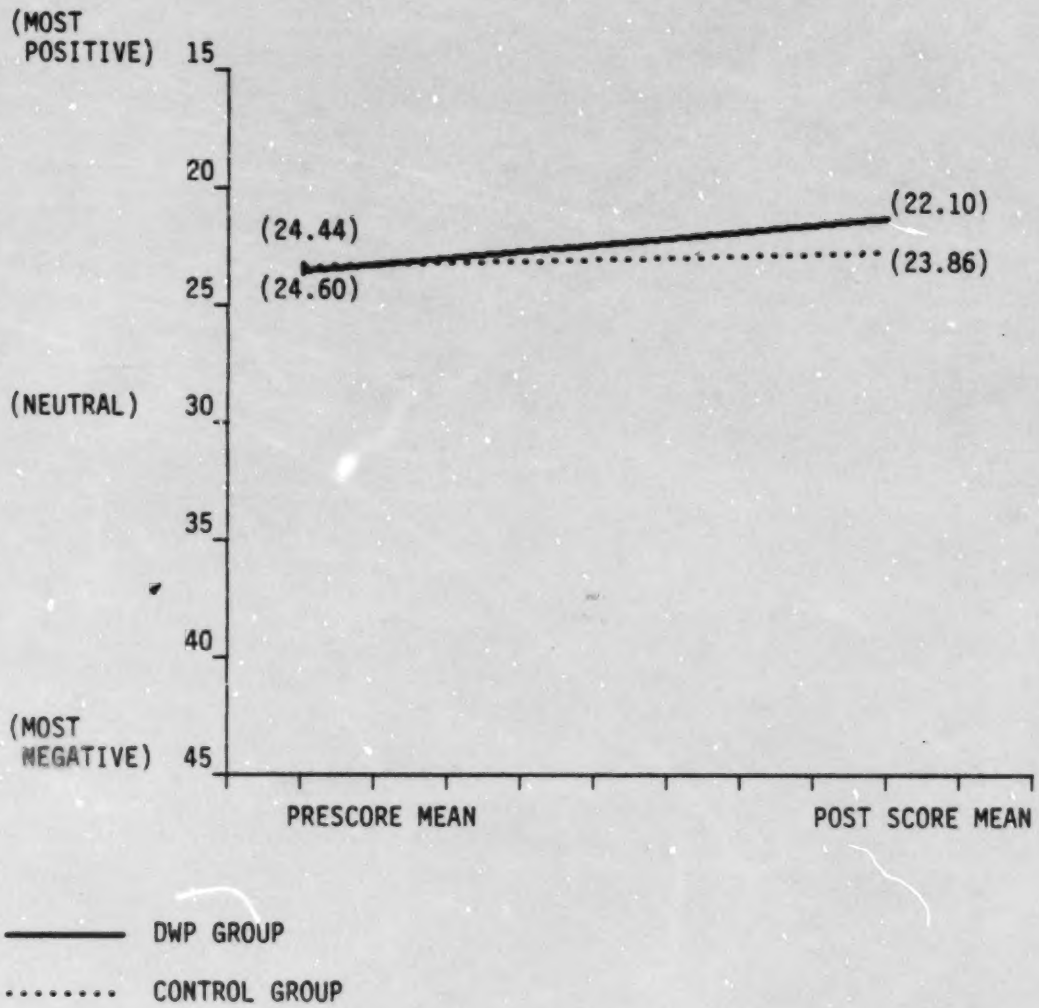
	-------- Attitude Survey --------	
	Pre Score* Mean	Post Score* Mean
DWP Group N = 488 (Pre) N = 454 (Post)	24.44	22.10[a]
Control Group N = 457 (Pre) N = 469 (Post	24.60	23.86[b]

* Most positive score = 15; most negative score = 45

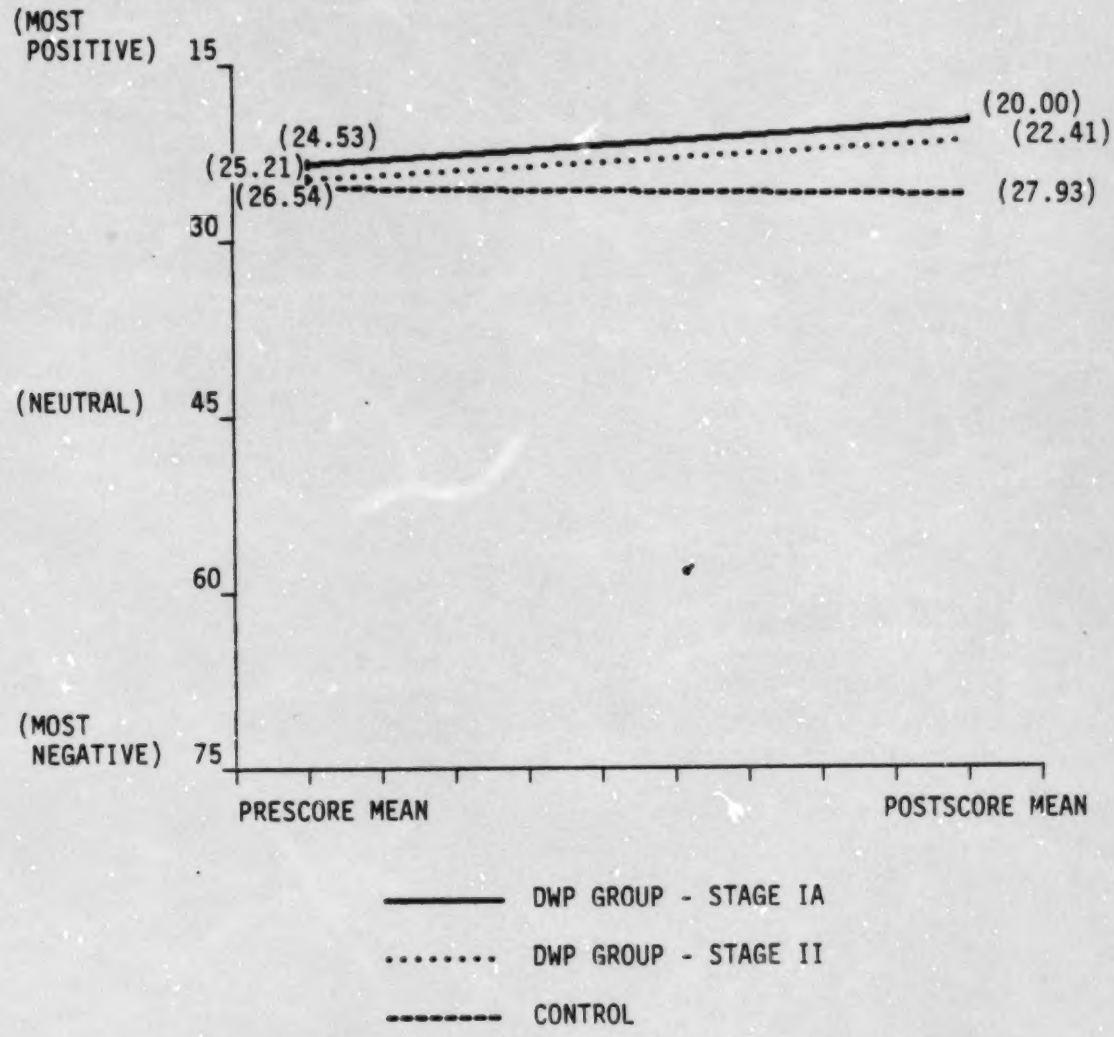
[a][b] These groups are significantly different from each other p .000

Figure 2

Developmental Writing Program and Control Group
Pre and Post Mean Scores on Student Attitude Survey



Groups' Pre to Post Mean Scores on
Toward Teaching Writing



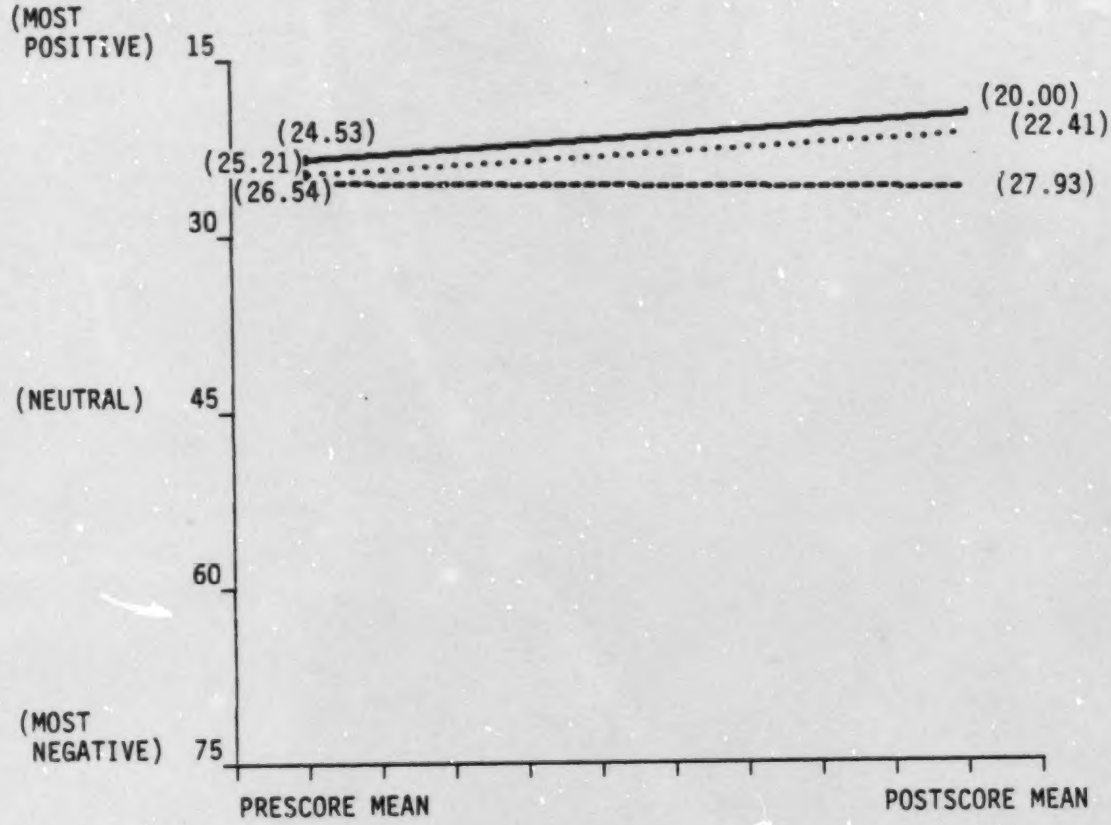
To answer the evaluation question:

> **Are teachers' instructional behaviors changed as a result of training received in the DWP?**

an interview procedure was employed. To measure changes in teacher behaviors in the teaching of writing, the Concerns-Based Adoption Mode (C-BAM) Levels of Use (LoU) focused-interview system was utilized with a sample of teachers in April 1984. According to researchers at the University of Texas (Hall, Loucks, Rutherford, Newlove, 1975), instructional change, or use of new instructional procedures, occurs at a different rate and in a different way for each individual. In order to assess that change in a systematic fashion they developed an eight-level interview procedure and assignment system. Levels range from no knowledge of an instructional procedure or program (0) to highly effective use of it (IVA, IVB) then, further, to searching for an even more effective alternative (VI). A 20 to 30 minute individual teacher interview is conducted, then a level assignment is made by a trained professional certified in the LoU system. Appendix J shows the LoU chart which includes all assignment levels and a description of typical behaviors for each level.

Because the DWP is a unique program model it was assumed by evaluators and program developers that, prior to inservice training in program use, teachers were not familiar with the DWP and did not teach writing in the DWP manner; thus all teachers were initially classified as non-users (level "O" on the LoU scale). It was anticipated, after 18 hours of training (see Appendix K) and either 14 (Stage I) or 8 (Stage IA and II) months of instruction, most teachers

Groups' Pre to Post Mean Scores on
Toward Teaching Writing



To answer the evaluation question:

> **Is teacher attitude toward writing different after DWP participation?**

two analyses were conducted. First, within-group average pre and post attitude survey scores were computed and compared for three treatment groups (all except Stage I). Second, average postscores were compared among these three groups. Table 10 shows results of both procedures. Based on results of a one-way ANOVA, the three groups of teachers were not significantly different in their attitudes toward writing at the beginning of the 1983-84 school year (column 1). At the end of the year, however, the Stage II DWP group exhibited significantly ($p \leq .05$) more positive attitudes toward writing than the Control Group (column 2). Teachers in the Stage IA DWP group were not substantially different in their attitudes toward writing than the other two groups. A score close to 30 indicates a positive attitude toward one's own writing and the teaching of writing; a score close to 150 indicates a negative attitude toward one's own writing and the teaching of writing.

Also shown in Table 10 are the three groups' scores for the two components of the attitude survey. Half of the survey questions related to a teacher's attitude toward his/her own writing. The other half of the survey included items related to attitude toward the teaching of writing. Groups did not significantly differ on the pre or post surveys in attitudes toward their own writing, nor were they significantly different on the pre survey in their attitudes toward teaching writing. On the post survey, however, both DWP groups were significantly more positive toward teaching writing than the Control Group (column 6).

Table 10

Two Developmental Writing Program and Control Groups' Pre to Post Mean Scores on Teacher Attitude Survey -- Total Score and Two Partial Scores

	Total Score		Attitude Toward Own Writing		Attitude Toward Teaching Writing	
	Pre Score Mean*	Post Score Mean*	Pre Score Mean**	Post Score Mean**	Pre Score Mean**	Post Score Mean**
DWP Group - Stage IA N = 33 (Pre) N = 27 (Post)	59.82	53.56[a][b]	35.08	31.15	25.21	22.41[a]
DWP Group - Stage II N = 45 (Pre) N = 33 (Post)	58.48	49.82[b]	33.80	29.82	24.53	20.00[a]
Control Group N = 13 (Pre) N = 16 (Post)	61.00	61.06[a]	34.46	33.12	26.54	27.94[b]

* most positive score = 30; most negative = 150

Groups with different letters are significantly from each other p .05.

** most positive = 15; most negative = 75

[a] These groups are not significantly different.

[b] This group is significantly different from the others p .05.

Figures 3 - 5 show pre and post mean scores and growth slopes for each group. As shown in figure 3, the control group experienced almost no growth from pre to post in total attitude score. Teachers in the two DWP groups exhibited significantly more positive attitudes about writing and writing instruction at the end of the year than at the beginning of the year. As show in figure 4, DWP teachers and Control Group teachers moved in a positive direction in their attitudes toward their own writing. In contrast, figure 5 indicates a more positive attitude toward the teaching of writing by the two DWP groups, but a less positive attitude by the Control Group at the end of the academic year than at the beginning of the year.

DWP teachers in Stage I were not included in this investigation because they had already experienced one-half year of DWP instruction. Evaluators and program developers assumed that this group would have substantially different attitudes about writing and writing instruction than the other three groups.

Figure 3

Two Developmental Writing Program and Control Groups' Pre to Post Mean Scores on Teacher Attitude Survey Total Score

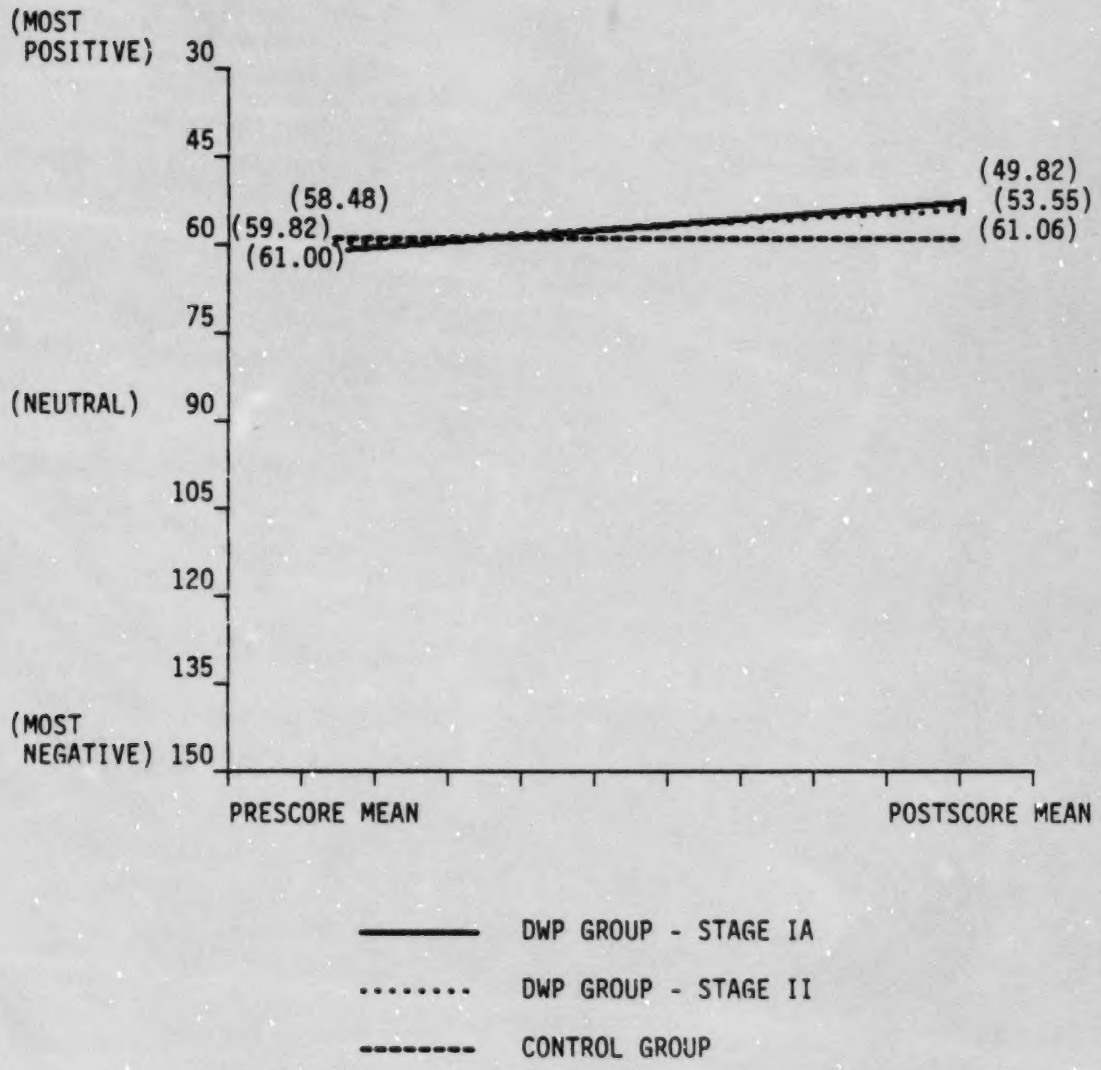


Figure 4

Two Developmental Writing Program and Control Group's Pre to Post mean Scores on Teacher Attitude Survey -- Attitude Toward Own Writing

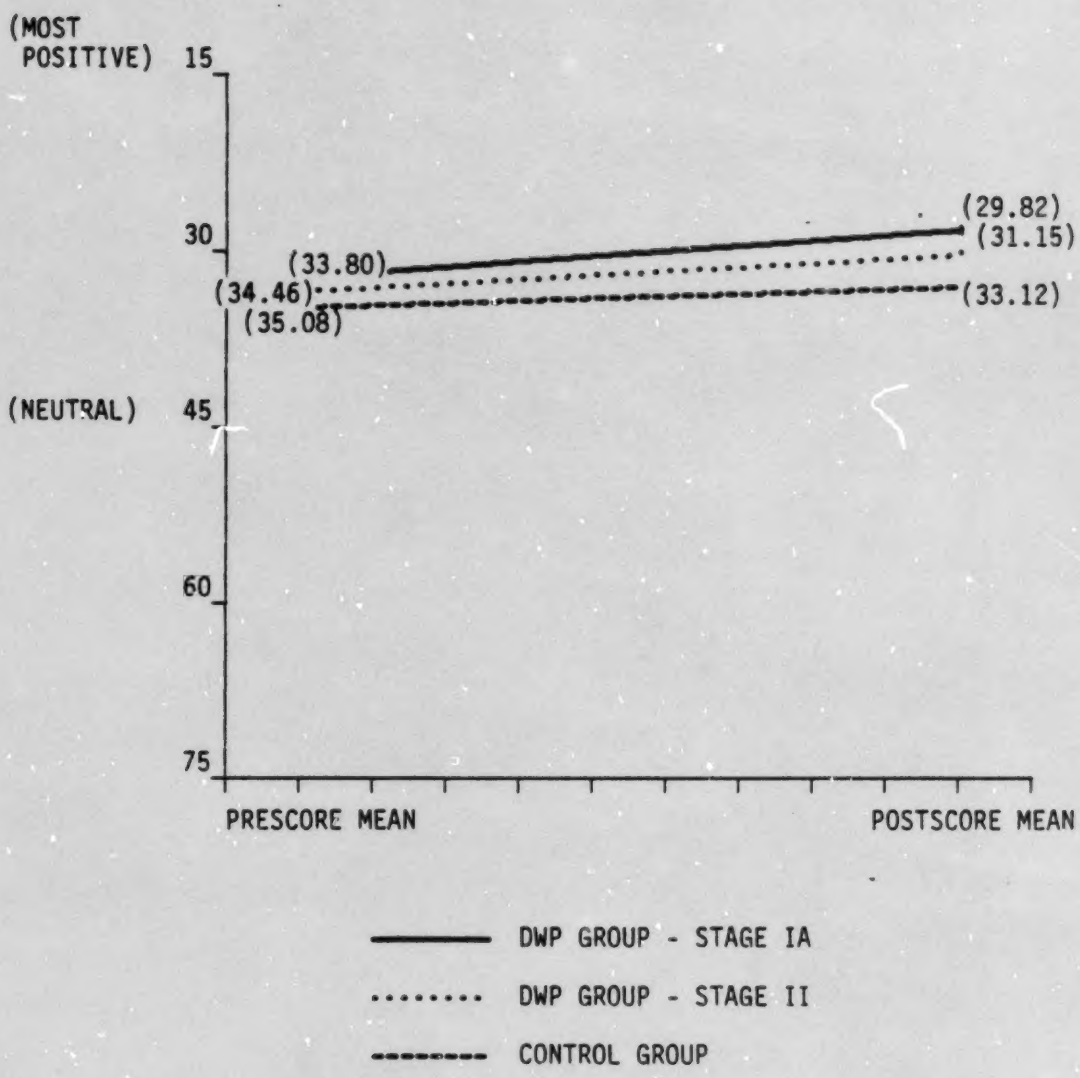
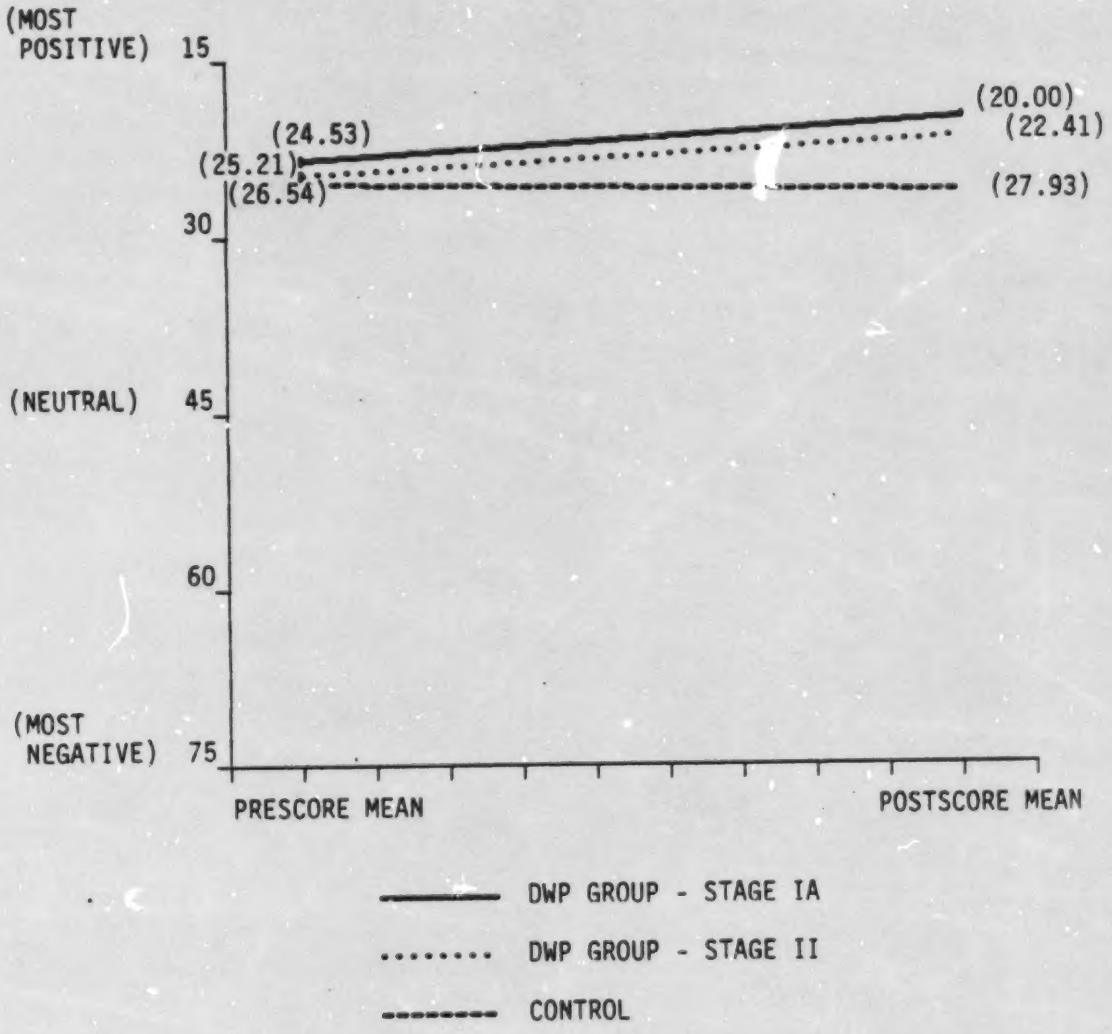


Figure 5

Two Developmental Writing Program and Control Groups' Pre to Post Mean Scores on Teacher Attitude Survey -- Attitude Toward Teaching Writing



To answer the evaluation question:

> Are teachers' instructional behaviors changed as a result of training received in the DWP?

an interview procedure was employed. To measure changes in teacher behaviors in the teaching of writing, the Concerns-Based Adoption Mode (C-BAM) Levels of Use (LoU) focused-interview system was utilized with a sample of teachers in April 1984. According to researchers at the University of Texas (Hall, Loucks, Rutherford, Newlove, 1975), instructional change, or use of new instructional procedures, occurs at a different rate and in a different way for each individual. In order to assess that change in a systematic fashion they developed an eight-level interview procedure and assignment system. Levels range from no knowledge of an instructional procedure or program (0) to highly effective use of it (IVA, IVB) then, further, to searching for an even more effective alternative (VI). A 20 to 30 minute individual teacher interview is conducted, then a level assignment is made by a trained professional certified in the LoU system. Appendix J shows the LoU chart which includes all assignment levels and a description of typical behaviors for each level.

Because the DWP is a unique program model it was assumed by evaluators and program developers that, prior to inservice training in program use, teachers were not familiar with the DWP and did not teach writing in the DWP manner; thus all teachers were initially classified as non-users (level "O" on the LoU scale). It was anticipated, after 18 hours of training (see Appendix K) and either 14 (Stage I) or 8 (Stage IA and II) months of instruction, most teachers

would use the DWP in at least a mechanical fashion (level III) and some teachers would experience routine, comfortable use of the DWP (level IVA).

Results of the LoU assignments of 30 teachers, interviewed in April 1984 are shown in table 11. As expected, most teachers (64%) in the sample were classified as "mechanical users"; a smaller percentage (30%) were "routine users." Two unexpected occurrences were: one teacher had not yet implemented the program (an intern was teaching the class), but was preparing to do so, and one teacher had refined and/or adapted the DWP to a particular group of science students.

Table 11

Developmental Writing Program
Levels of Use Assignments -- April 1984

	Number of DWP teachers who scored at this level
0 Non-Use	0
I Orientation	0
II Preparation	1 (3% of sample)
III Mechanical Use	19 (64% of sample)
IVA Routine Use	9 (30% of sample)
IVB Refinement	1 (3% of sample)
V Collaboration	0
VI Refocusing	0
TOTAL	30 (100% of sample)

To further investigate changes in teacher behavior, the sample was separated into the three teacher groups: Stages I, IA, II. Table 11 shows the frequency of assignment for each level by group. Second year DWP teachers (Stage I) appear to cluster in levels III and IVA; most first year DWP teachers with access to a school-based resource (Stage IA) fell into level III; first year DWP teachers with no access to a school-based resource (Stage II clustered in levels III and IVA. It is important to note that a higher percentage of Stage II teachers moved into the "routine user" category than Stage IA teachers.

Thus, it does not appear as if having an experienced DWP teacher on-campus facilitated routine implementation of the program or sped up the change process. One explanation for this surprise occurrence may be that first-year teachers did not communicate with each other about the DWP. Indeed, as stated in section 4.1, most Stage IA teachers did not mention asking for or receiving assistance from their Stage I colleagues.

Table 12

Developmental Writing Program
Levels of Use Assignments by Teacher Group -- April 1984

	STAGE I (2nd year DWP teachers)	STAGE IA (1st year DWP teachers - access to school-based resource)	STAGE II (1st year DWP teachers - no access to school-based resource)
0 Non-Use	---	---	---
I Orientation	---	---	---
II Preparation	---	---	7%
III Mechanical Use	57%	89%	50%
IVA Routine Use	28%	11%	43%
IVB Refinement	14%	---	---
V Collaboration	---	---	---
VI Refocusing	---	---	---
Number in sample	7	9	14

5.0 CONCLUSIONS AND RECOMMENDATIONS

This section includes conclusions about the 1983-84 Developmental Writing Program and recommendations for its continuation. Conclusions are based on data reported in section 4.0 of this document, are presented first. One overall program statement is followed by results for the specific evaluation questions. Recommendations, presented last, resulted from synthesis of all data and conversations with program developers.

5.1 Conclusions

Overall. The Developmental Writing Program, which operated in 35 schools in grades 4-6, involved 98 teachers and approximately 3,000 students during the 1983-84 school year. The program appeared to be implemented as intended. Its implementation resulted in improved teacher and student attitudes toward writing and improved student writing. Most teachers changed their instructional behaviors as a result of the Developmental Writing Program.

Implementation. Almost all participating teachers felt they were adequately trained to teach the Developmental Writing Program. Although sufficient instructional time and materials were cited as program weaknesses, most teachers instructed students for at least one and one-half to two hours per week in writing skills using the Developmental Writing Program approach. Almost all teachers were enthusiastic about the Developmental Writing Program, and students reportedly enjoyed and benefitted from this instruction.

Students' Pre to Post Writing. At least 70 percent of students who participated in the Developmental Writing Program exhibited improved writing skills after eight months of Developmental Writing Program instruction. Fifty-seven (57) percent of students who did not participate in the Developmental Writing Program, but who may have

received other types of writing instruction, also exhibited improved writing skills after eight months in school. However, when compared with these other instructional methods, the Developmental Writing Program was significantly more effective in improving students' writing skills.

Students who were instructed by first-year Developmental Writing Program teachers made about the same progress in writing as those instructed by second-year Developmental Writing Program teachers.

Students' Achievement Test Scores. The Developmental Writing Program was at least as effective as more traditional types of language arts instruction in teaching students basic language skills. In some instances, Developmental Writing Program instruction appeared more effective than traditional instructional methods (e.g., language books, worksheets). Fourth, fifth, and sixth grade students who participated in the Developmental Writing Program scored at least as well as a comparable group of non-participants in reading comprehension, language mechanics, and language expression.

Students' Attitudes Toward Writing. Students who participated in the Developmental Writing Program and in other types of writing instruction had more positive attitudes toward writing at the end of the school year than at the beginning of the year. However, Developmental Writing Program participants were significantly more positive in their attitudes toward writing than the comparable group of non-participants.

Teachers' Attitudes Toward Writing. One group of first-year Developmental Writing Program teachers were significantly more positive in their overall attitudes toward writing at the end of the school year than a comparable group of non-participating teachers. All first-year Developmental Writing Program teachers had significantly more positive attitudes toward the teaching of writing than teachers in the non-participating group.

Teachers' Instructional Behaviors. Almost all participating Developmental Writing Program teachers changed their instructional behavior in the area of writing. More than half the total group of first and second-year teachers implemented the program in a "mechanical" fashion. That is, most effort focused on managing instructional time, organizing materials, finding topics, determining methods for editing and publishing, and learning about the requirements of the program. About a third of Developmental Writing Program teachers implemented the program on a "routine" basis. That is, they were comfortable and consistent in their approach to the program. For these routine users, time, materials, topics, editing, publishing and other routines were clearly established; writing classes were planned and operated with a minimum amount of stress. These percentages are as expected during the initial phases of a new instructional program.

5.2 Recommendations

1. The Developmental Writing Program should be maintained in present locations and expanded, over the next three years, to include all elementary schools which house grades 4, 5, and/or 6. To alleviate concerns about time constraints, instructional time for the Developmental Writing Program should be specified in the district's "redefined day" plan.

2. With the exception of reorganization of the introductory portion of the training workshops, teacher training sessions for "new" Developmental Writing Program teachers should be conducted in the same manner as 1983-84 training sessions. Follow-up sessions for "old" Developmental Writing Program teachers should include topics such as: evaluation/grading of students' writing, using computers in writing instruction, outside publishing of student writing, and enhancing one's own writing skills. Training activities for experienced teachers should take into account the mechanical or routine status of their implementation.

3. More detailed program materials should be developed and provided to both old and new Developmental Writing Program teachers. These materials should specifically emphasize topic ideas and prewriting activities.

4. Increased communication/collaboration among school-based Developmental Writing Program colleagues should be encouraged.

5. In order to assure continuity in writing instruction and consistent student progress in writing, program developers should communicate with secondary supervisors and teachers, particularly those at grades 7 through 9, to determine the extent of writing instruction at these levels.

6. Continued evaluation of the Developmental Writing Program is recommended. The third year's evaluation should focus on investigations of student progress in writing by grade level and ability level. Based on these longitudinal investigations, expectations for students can be set.

References

Emig, J. (1971) The composing processes of twelfth graders (Research Report No. 13). Urbana, Illinois: National Council of Teachers of English, Committee on Research. (ERIC Document Reproduction Service No. ED 038 252)

Graves, D. H. (1974) Children's writing: Research directions and hypotheses based upon an examination of the writing process of seven-year-old children (Doctoral dissertation, State University of New York at Buffalo, 1973). Dissertation Abstracts International, 34, 6255A. (University Microfilms No. 74-08375) (ERIC Document Reproduction Service No. ED 095 586)

Hall, Gene E, Loucks, Susan F., Rutherford, William L., & Newlove, Beulah W. (1975). Levels of use of the innovation: A framework for analyzing innovation adoption. Journal of Teacher Education, 26(1).

Murray, D. (1968) A teacher teaches writing. Boston: Houghton-Mifflin Co.

APPENDIX A

Developmental Writing Program

Control Group End-of-the-Year Survey

Hillsborough County Public Schools
Developmental Writing Program Evaluation
1983-84

PLEASE COMPLETE THIS BRIEF SURVEY AND RETURN IT WITH YOUR STUDENTS' MATERIALS.

1. Do you teach your students how to write?

   ____ Yes
   ____ No

2. How often do you teach writing to your students? (may include poetry, creative writing, writing research papers, etc.)

   _____ times per week

3. About how long do these lessons usually last?

   _____ minutes per lesson

4. What materials do you use to teach writing? (list names of books, filmstrips, tapes, software, etc.)

   _____
   _____
   _____
   _____

5. What activities do you use and/or assign when teaching writing?

   _____
   _____
   _____
   _____

6. Are you satisfied with your present writing program? (Circle one.)

   Very Satisfied Satisfied Neutral Dissatisfied Very Dissatisfied

THANKS!

BEST COPY AVAILABLE

APPENDIX B

Developmental Writing Program

Teacher Interview*

* Adopted from the LoU Interview, University of Texas CBAM Project

Date_____

Teacher Stage I IA II

Cycle 2 3

Hillsborough County Public Schools
Developmental Writing Program Evaluation
Teacher Interview*

1-2-3-4	Are you currently using the Developmental Writing Program?
	If yes, turn page. If no, continue.
	NO
	Have you ever used it in the past? If so when? Why did you stop?
	If yes, go to PAST USERS (Below)
	If no, continue.
1/2	Have you made a decision to use the Developmental Writing Program in the future?
1/2	If so, when will you begin?
Knowledge	Can you describe the Developmental Writing Program to me as you see it?
Acquiring Information	Are you currently looking for any information about the Developmental Writing Program? What kinds? For what purposes?
Knowledge	What do you see as the strengths and weaknesses of the Developmental Writing Program? Give examples if necessary.
Assessing	At this point in time, what kinds of questions are you asking about the Developmental Writing Program? Give examples if necessary.
Sharing	Do you ever talk with others and share information about the Developmental Writing Program? What do you share?
Planning	What are you planning with respect to the Developmental Writing Program? Can you tell me about any preparation or plans you have been making for the use of the Developmental Writing Program?
Concern	Do you have any concerns about using the Developmental Writing Program in your setting?

PAST USERS

Can you describe for me how you organized your use of the Developmental Writing Program, what problems you found, what its effects appeared to be on students?

When you assess the Developmental Writing Program at this point in time, what do you see as the strengths and weaknesses?

NOW, GO TO ABOVE SECTION, STARTING WITH QUESTION MARKED 1/2.

*Adapted from the LOU Interview, University of Texas CBAM Project

(11/83 and 3/84)

YES

Open-Ended	Please describe for me how you use the Developmental Writing Program. (Ask sufficient questions to cover minimal criteria for use.)
Assessing/ Knowledge	What do you see as the strengths and weaknesses of the Developmental Writing Program in your situation? (Have you made any attempt to do anything about weaknesses? Probe those they mention specifically.)
Acquiring Information	Are you currently looking for any information about the Developmental Writing Program? What kind? For what purposes?
Level V	Do you work with others in your use of the Developmental Writing Program? Do you meet on a regular basis? Have you made any changes in your use of the Developmental Writing Program based on this coordination? How about other DWP teachers at your school?
	If yes, go to Level V Probes (Below)
Sharing	Do you ever talk with others about the Developmental Writing Program? What do you tell them?
Assessing	(Have you considered any alternatives or different ways of doing things with the program?) Are you doing any evaluating, either formally or informally, that would affect your use of the Developmental Writing Program? Have you received any feedback from students that would affect the way you're using the Developmental Writing Program? What have you done with the information you got?
III/IVA/ IVB	Have you made any changes recently in how you use the Developmental Writing Program? What? Why? How recently? Are you considering making any changes?
Planning/ Status Reporting	As you look ahead to later this year, what plans do you have in relation to your use of the Developmental Writing Program?
Concerns	Do you have any concerns about using the Developmental Writing Program in your setting?

LEVEL V PROBES

1. Please describe for me how you work together. (What things do you share with each other?)
2. What do you see as the effects of this collaboration?
3. Are you looking for any particular kind of information in relation to this collaboration?
4. Do you talk with others about your collaboration? If so, what do you share with them?
5. Have you done any formal or informal evaluation of how your collaboration is working?
6. What plans do you have for this effort in the future?

If you have enough evidence to place the person at Level V, ask final question here.

If you do not think the person is a Level V, go to Question Sharing.

(11/83 and 3/84)

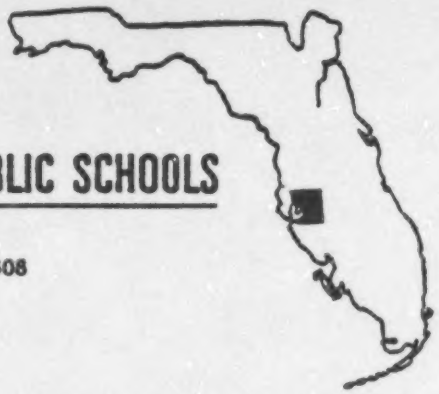
APPENDIX C

Developmental Writing Program

Topics and Procedures

HILLSBOROUGH COUNTY PUBLIC SCHOOLS

P.O. BOX 3408
TAMPA, FLORIDA 33601-3408



SCHOOL BOARD
JOE E. NEWSOME, Chairman
ROLAND H. LEWIS, Vice Chairman
CECILE W. ESSRIG
A. LEON LOWRY
R. SONNY PALOMINO
SAM RAMPELLO
MARION S. RODGERS

Superintendent of Schools
RAYMOND O. SHELTON

School Administrative Center
901 EAST KENNEDY BLVD. (TAMPA)
(813) 272-4000
SUNCOM 571-4000

MEMORANDUM

TO: Participating DWP Teachers and Control Group Teachers

FROM: Susan D. Turner, Supervisor of Evaluation
Department of Testing and Evaluation

DATE: April 30, 1984

SUBJECT: Post-Data Collection for the Developmental Writing Project

As part of the program evaluation of the Developmental Writing Project (DWP), you are asked to collect post-data from your language arts or DWP students. This memo gives you directions for collecting and returning two types of information, Student Attitude Surveys and Writing Samples.

STUDENT ATTITUDE SURVEYS

1. Have students complete the attitude survey the week of May 14-18. It should be administered to all students in your language arts or developmental writing class. Each student will need a survey and a standard answer sheet.

2. For each student, the following information must be written with pencil on the standard answer sheets: school, teacher, grade and name. The school number and student number may be bubbled in by the teacher or the student; please be sure it is correct.

3. Read the instructions with the students. Be sure they understand the three ratings. Any words students do not know may be read to them. In fact, you may read all the items to the students if you wish.

4. Do not staple anything to the standard answer sheet.

WRITING SAMPLES

1. A writing production sample will be part of the program evaluation of the developmental writing program. Students wrote two pre-program samples and will write two post-program samples. The two post-samples will be collected during the weeks of May 14-18 and May 21-25.

2. Students will be assigned a topic each time they write. The same two topics are used in the pre/post samples.

3. All student writing must be done in pencil on the forms provided.

4. On the writing day, tell the students the following:

> Today you are going to do some writing. The reason you are writing now is to compare an end-of-the-year writing sample with a beginning-of-the-year writing sample to see if you have improved. You may remember that we wrote on these same two topics way back in September.

5. Distribute writing forms. Read the directions aloud while students read them silently.

6. After reading the directions, tell students:

> You have 30 minutes in which to write. Please write or print clearly. Make any revisions by crossing out neatly.

7. If students request to do so, they may use scratch paper.

8. If students ask to use the dictionary or for teacher help, tell them they are to do their best on their own.

9. If students ask how long their writing should be, tell them to try to keep it to one page. If their writing must be longer, have them complete it on a sheet of notebook paper with their name in the upper right-hand corner. Staple the second page to the first.

10. After ten minutes, tell students that if they are still thinking, they should begin writing now.

11. After 25 minutes, tell students they have five minutes to bring their papers to a close.

12. Tell students that if they finish early, they are not to leave their seats. Direct them to read a library book or to some other seat activity.

13. Collect all materials. Check all writing samples to make sure there is a legible first and last name on each. Alphabetize each set of writing samples.

14. Repeat the procedures for the second writing sample with the following introduction:

> Last week you did some writing as part of an evaluation of your writing skills. Today you are going to write some more as part of the same project. This will allow us to select your best writing sample of the two.

Please try to get attitude surveys and two writing samples from all students, including those who are absent on the day the class writes. However, do not collect any samples after May 25.

Also, please include an updated alphabetical class roster along with the surveys and writing samples.

Return all materials no later than May 23 to Susan Turner, Department of Testing and Evaluation, School Administrative Center. If you have any questions or concerns, feel free to contact Susan Turner at 272-4344.

PLEASE NOTE: We are aware that the manner in which these samples are collected is not congruent with the way students usually write (i.e., pre-writing, drafting, editing, final copy writing). However, a standardized writing assessment procedure is necessary for this evaluation. Please follow the directions as stated.

Thank you for your cooperation.

SDT/dsr
Enclosures

HILLSBOROUGH COUNTY PUBLIC SCHOOLS

Writing Production Sample A

Name_____ School_____ Date_____

Teacher_____ Grade_____

DIRECTIONS: Pets are very special to many people. Write about a pet you have or would like to have.

HILLSBOROUGH COUNTY PUBLIC SCHOOLS

Writing Production Sample B

Name_____ School_____ Date_____

Teacher_____ Grade_____

DIRECTIONS: People are frightened by different things. Think about what scares you and write about it.

APPENDIX D

Developmental Writing Program

Teacher (30-items) and Student (15-items) Attitude Surveys

A = "I strongly agree"
B = "I agree"
C = "I cannot decide"
D = "I disagree"
E = "I strongly disagree"

1. I enjoy sharing my writing efforts with others.
2. I haven't written anything of substance in years.
3. I like learning and using new words.
4. I cannot express how I feel in my writing.
5. I like to do different kinds of writing.
6. Creative writing has never interested me.
7. I can clarify my ideas when I write.
8. Writing is a chore.
9. I am embarrassed by my writing style.
10. I don't really have much to say when I write.
11. I like to see my thoughts written down.
12. Sometimes I get excited about a topic when I am writing.
13. I don't write because my vocabulary is not broad enough.
14. My writing is usually boring.
15. I do not like other people to read my writing.
16. Writing is a very important skill.
17. Teaching writing gets to be boring after awhile.
18. Students seldom have valuable ideas to contribute in a pre-writing discussion.
19. Time spent on writing instruction transfers to other areas of language arts.
20. Good writing can be at least as effective as a picture.
21. I've had poor results teaching students to write.
22. Preparing students to write is not as important as the actual writing.
23. Students should write at least several times a week.
24. All other things being equal, an enthusiastic teacher will motivate students more than an unenthusiastic one.
25. There are few good strategies to use in teaching writing.
26. Other people's writing gives me ideas to improve my own writing.
27. All students can improve their writing skills.
28. I encourage students to revise their writing.
29. I'd like to teach my students to write well.
30. Students should do many different types of writing.

A = "I agree"
B = "I cannot decide"
C = "I disagree"

1. There are many interesting things I can write about.
2. I like to do lots of different kinds of writing.
3. I feel proud when the teacher displays my writing in the classroom.
4. I like using new words.
5. I don't like my friends to read what I have written.
6. I like to write.
7. It's fun to write about an interesting topic.
8. I can never think of anything to write.
9. I like to tell how I feel about things in writing.
10. I don't think that I write very well.
11. I only write when I have to.
12. I think sometime I might write a book.
13. I can learn to be a better writer.
14. I like to read or listen to what my classmates write.
15. I don't like to make up stories and write them down.

APPENDIX E

Developmental Writing Program

Interim Report -- November 1983

Summary of Developmental Writing Program Classroom Observations and Teacher Interviews -- November 1983

During November 1983, Susan Turner and Lore Nielsen conducted interviews and observations in a random sample of 23 Developmental Writing Program teachers' classrooms; equivalent proportions of teachers were selected from each of the three stages of the program. Purposes of these activities included: gathering attitudinal information from teachers and students about the Developmental Writing Program, and observing how teachers actually implemented the program in their classrooms. After all interviews and observations were completed information was compiled and summarized and conclusions about program operation were shared with Jackie Matthews and George Mitchell. The conclusions that we shared were as follows:

.In the 23 classrooms we observed, all of you did some kind of writing lesson for us. Usually the lesson consisted of pre-writing and drafting. Sometimes we saw editing and publishing occur. Once or twice we observed students writing in their journals. The chalkboard or chart paper and student pencil and paper were the most often used materials. Sometimes a lesson included use of the overhead projector and/or pictures. In one or two instances, some worksheets from the Easy Writer book were used. Average lengths of Developmental Writing Program lessons are shown in the chart on the following page. As you can see, they all varied in length and there were some differences noted among the three stages of teachers.

	LENGTH OF DWP OBSERVED LESSONS		
	Stage I	Stage IA	Stage II
Total Lesson	Range 25-50 minutes Average 41 minutes	Range 15-60 minutes Average 36 minutes	Range 35-65 minutes Average 46 minutes
Actual Student Writing Time	Range 7-33 minutes Average 17 minutes	Range 10-45 minutes Average 20 minutes	Range 12-45 minutes Average 23 minutes

We defined an actual lesson length as everything (start to finish) from pre-writing to when the whole writing activity or task seemed to be over. In order to find out how long students actually wrote within each lesson, we recorded the time that most students began to put pencil to paper and the time that most students stopped writing. We were pretty lenient with these student-writing times and usually counted editing situations within this time frame.

.Students responded, on the whole, very positively when asked about the Developmental Writing Program and their involvement in it. The table on the next page summarizes student responses for the 22 classrooms in which students were interviewed. In most cases, student responses closely corresponded to classroom observational information and teacher interview data.

.In almost all of the 23 classrooms, we found evidence that writing occurred on a regular basis. Student writing products and various writing procedure directives were posted in 20 classrooms. Student folders and resource materials (e.g., word banks, dictionaries, and filmstrip kits) were obvious in 22 classrooms. Writing centers were noted in 14 classrooms.

.Of particular interest to Jackie and George were the various stages of implementation in which people found themselves. After interviewing 23 teachers and about 50 students and observing 23 lessons, we noted that:

1. About three people don't use the DWP regularly as it is intended, but do use a few of the parts or ideas from the DWP.
2. One person is not yet using the DWP but is preparing to do so and has set a date to begin.
3. About 12 people are using the DWP including most or all of its parts, but are having some concerns about time and management.

Summary of DWP Student Responses

TYPE OF QUESTION	YES	NO
Do you get to write or have a writing lesson everyday?	14	8
When you have a writing lesson do you first talk about what you're going to write?	21	1
Do you write drafts of things?	18	3
Do kids in this class work together in editing each others' writing?	15	7
To help in choosing words for writing, have you started using the thesaurus?	12	10
dictionary?	21	1
other word books?	14	8
Do you feel like you can use them easily?	21	0
Can you use them whenever you want?	20	1
Have you found some new words to use when you write?	20	1
Do you keep your writing in a personal folder?	22	0
Do you use your folder during writing class?	19	3
Is there a special place in your room that's called or used as a writing center?	12	10
Do you ever get to go there and do some writing?	7	5
Do other kids get to go to the writing center?	8	4
Have you had a chance to show off or put up anything you've written yet?	22	0
Do you read what you write to other people?	19	3
Has (principal) been in to see or hear anything you or other kids have written?	16	6

What do you like best about this whole writing thing?

- .You get to use your imagination.
- .I think all new kids to this school should do it.
- .I like the journals; they're secret.
- .discovering/learning new words
- .word webbing (other specific activities were mentioned by a number of students)
- .I like to make up stories.
- .Just writing is fun.
- .You learn a lot.
- .We get to express what we feel.
- .I like the substitutions.
- .writing poems
- .I like the part where you keep it for a year, then go back and read it.
- .Writing is beautiful.
- .The teacher shows you how to write better.
- .It gets you away from routine work.
- .You get to work individually.
- .You learn new things you can use in your other work, such as proof-reading.

4. About six people are using the DWP on a pretty routine basis and seem comfortable with most everything about it.
5. One person has refined the DWP and modified it for his/her own unique situation.

Those realities seem to be in line with reasonable expectations for a new program like the Developmental Writing Program. You can see that this type of information is helpful to the program developers in knowing what types of training sessions to plan and what kinds of visitations to make.

Lore and I thank you for your cooperation during our visits. The next observation/interview cycle will take place in March with a different sample group. Those of you who will be involved will be informed in plenty of time.

APPENDIX F

Developmental Writing Program

Focused Holistic Score Scale

Focused Holistic Score Scale

A five-point focused holistic scale will be used in this scoring activity. Only four of the scale points, 1 through 4, will be used to judge the quality of performance, however. The zero value will be reserved for certain anomalous responses and it cannot be used unless permission has been given by the trainer. Each scale value is defined by an anchor statement and by several actual writing samples which are judged to illustrate the specified level of performance. The scale points and their respective anchor statements are given below. Each is restated in the following chapter, where Anchor Papers and rational statements are provided to further explicate each scale point.

4 = The composition is clearly focused and the subject-matter is well-developed. The writer's delivery is well-organized, logical, and coherent. Sentences are correctly formed, and the writer has varied sentence structure and length. The writer's language is effective and is appropriate for standard, informal written discourse. The writer shows evidence of a consistent command of written language mechanics.

3 = The composition is reasonably focused and the subject-matter is adequately developed. The writer's delivery is sufficiently organized, logical, and coherent. Most sentences are correctly formed, and the writer's general command of sentence formation is evident. The writer has a reasonable variety of sentence structure and length. Overall, the writer's language is reasonably effective and is acceptable for standard, informal written discourse, though minor deviations may exist. The writer shows evidence of an overall understanding of written language mechanics; deviations are minor.

2 = The composition's focus is weak. Although the subject-matter is discernible, development is minimal. Organization, logic, and coherence are minimal, but present. Some understanding of sentence formation is evident, but deviations may predominate. If the writer has a variety of sentence structure and length, it may be limited. Ineffective and inappropriate language use may cause the reader to question the writer's understanding of standard, informal written discourse. Some knowledge of written language mechanics is evident, but may be weak.

1 = The composition's focus is very weak, if present. The subject-matter may be marginally discernible, but is undeveloped. The composition is marred by weaknesses in organization, logic, and coherent delivery. Sentence formation errors are obtrusive, and the composition has little if any variety in sentence structure and length. Deviations from language expected in standard, informal written discourse are major, and command is not evident. The writer shows little if any evidence of an understanding of written language mechanics.

0 = The composition cannot be scored. YOU MUST RECEIVE THE TRAINER'S PERMISSION TO ASSIGN A "0".

APPENDIX G

Developmental Writing Program

List of Raters

List of Developmental Writing Program Raters

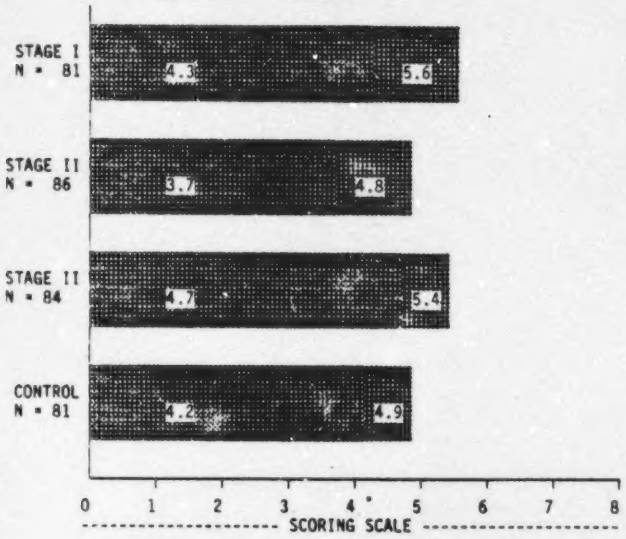
Rater	School
Sandy Coleman	Bellamy Elementary School
Helen Culver	Belle Witter Elementary School
Rita Dugan	Brooker Elementary School
Debra Larry	Palm River Elementary School
Cathy Messina	Chiaramonte Elementary School
Shirley Robinson	Pinecrest Elementary School
Beverly Stenz	Belle Witter Elementary School
Doyle Timmons	Foster Elementary School
Dolores Vaglica	Belle Witter Elementary School
May Ann Valdez	Town & Country Elementary School

APPENDIX H

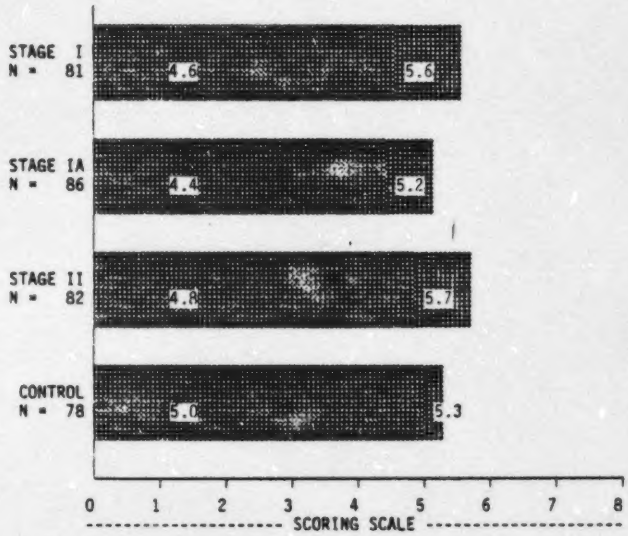
Writing Sample Growth Charts and
Pre-Post Differences by Grade

Writing Sample Growth Charts and Pre-Post Differences by Grade

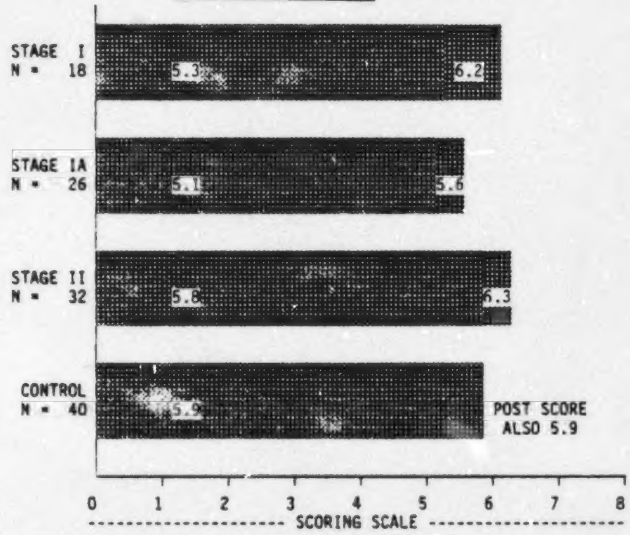
GRADE 4



GRADE 5



GRADE 6



AVERAGE SCORE ON PRE WRITING SAMPLES, AUGUST 1983

AVERAGE SCORE ON POST WRITING SAMPLES, MAY 1984
(GROWTH IS INDICATED BY LENGTH OF BAR)

Developmental Writing Program
Pre to Post Differences Among Treatment Groups

GRADE 4

Group	N	Number of students whose pre to post scores were: less than or equal to the average difference*	greater than the average difference*
Stage I	81	40 (50%)	41 (50%)
Stage IA	86	50 (58%)	36 (42%)
Stage II	84	59 (70%)	25 (30%)
Control	81	63 (78%)	18 (22%)

*Average pre to post difference across all groups = 1 point

GRADE 5

Group	N	Number of students whose pre to post scores were: less than or equal to the average difference*	greater than the average difference*
Stage I	81	48 (59%)	33[a] (41%)
Stage IA	86	56 (65%)	30[a] (35%)
Stage II	82	43 (52%)	39[a] (48%)
Control	78	68 (87%)	10[b] (13%)

* Average pre to post difference = 1 point
[a] These groups are not significantly different from each other.
[b] This group is significantly different from the others.

GRADE 6

Group	N	Number of students whose pre to post scores were: less than or equal to the average difference*	greater than the average difference*
Stage I	18	9 (50%)	9 (50%)
Stage IA	26	18 (69%)	8 (31%)
Stage II	32	16 (50%)	16 (50%)
Control	40	31 (77%)	9 (23%)

* Average pre to post difference across all groups = .5 points

Developmental Writing Program
Significance of Pre to Post Differences for Each Treatment Group on Writing Samples

GRADE 4

Group	N	Number (%) of students who experienced: Loss (-) from pre to post	No gain (0) from pre to post	Gain (+) from pre to post	Level of Significance
Stage I	81	5 (6%)	10 (12%)	66 (82%)	.00*
Stage IA	86	8 (9%)	15 (18%)	63 (73%)	.00*
Stage II	84	15 (18%)	12 (14%)	57 (68%)	.00*
Control	81	10 (12%)	11 (14%)	60 (74%)	.00*

* denotes statistical significance

GRADE 5

Group	N	Number (%) of students who experienced: Loss (-) from pre to post	No gain (0) from pre to post	Gain (+) from pre to post	Level of Significance
Stage I	81	12 (15%)	7 (9%)	62 (76%)	.00*
Stage IA	86	12 (14%)	17 (20%)	57 (66%)	.00*
Stage II	82	13 (16%)	5 (6%)	64 (68%)	.00*
Control	78	19 (24%)	18 (23%)	41 (53%)	.00*

* denotes statistical significance

GRADE 6

Group	N	Number (%) of students who experienced: Loss (-) from pre to post	No gain (0) from pre to post	Gain (+) from pre to post	Level of Significance
Stage I	18	6 (33%)	1 (6%)	11 (61%)	.03*
Stage IA	26	7 (27%)	4 (15%)	15 (58%)	.06
Stage II	32	7 (22%)	3 (9%)	22 (69%)	.00*
Control	40	15 (37%)	12 (30%)	13 (33%)	.95

* denotes statistical significance

APPENDIX I

Developmental Writing Program

Frequency Distributions

Frequency Distribution

Differences in Pre and Post Sum Scores

Group: Stage I (N = 180)

	2	3	4	5	6	7	8
Sum of 2 Pre-writing Samples	15	29	64	38	25	6	3
Sum of 2 Post-writing Samples	0	15	27	50	45	34	9

Below Average	Pre = 60% Post = 23%	Average and Above Average	Pre = 40% Post = 77%

Frequency Distribution

Differences in Pre and Post Sum Scores

Group: Stage IA (N = 198)

	2	3	4	5	6	7	8
Sum of 2 Pre-writing Samples	24	48	54	56	13	3	0
Sum of 2 Post-writing Samples	6	17	45	68	53	6	3

Below Average	Pre = 64% Post = 34%	Average and Above Average	Pre = 36% Post = 66%

Frequency Distribution

Differences in Pre and Post Sum Scores

Group: Stage II (N = 198)

	2	3	4	5	6	7	8
Sum of 2 Pre-writing Samples	3	22	59	75	28	10	1
Sum of 2 Post-writing Samples	3	6	24	61	71	28	5

Below Average	Pre = 42% Post = 17%	Average and Above Average	Pre = 58% Post = 83%

Frequency Distribution

Differences in Pre and Post Sum Scores

Group: Control (N = 199)

	2	3	4	5	6	7	8
Sum of 2 Pre-writing Samples	15	22	55	55	36	13	3
Sum of 2 Post-writing Samples	3	20	40	54	64	16	2

Below Average	Pre = 46% Post = 32%	Average and Above Average	Pre = 54% Post = 68%

APPENDIX J

LoU Chart

BEST COPY AVAILABLE

LEVELS OF USE

FIGURE I — LoU CHART

SCALE POINT DEFINITIONS OF THE LEVELS OF USE OF THE INNOVATION	CATEGORIES		
Levels of Use are distinct states that represent observably different types of behavior and patterns of innovation use as exhibited by individuals and groups. These levels characterize a user's development in acquiring new skills and varying use of the innovation. Each level encompasses a range of behaviors, but is limited by a set of identifiable Decision Points. For descriptive purposes, each level is defined by seven categories.	KNOWLEDGE That which the user knows about characteristics of the innovation, how to use it, and consequences of its use. This is cognitive knowledge related to using the innovation, not feelings or attitudes.	ACQUIRING INFORMATION Solicits information about the innovation in a variety of ways, including questioning resource persons, corresponding with resource agencies, reviewing printed materials, and making visits.	SHARING Discusses the innovation with others. Shares plans, ideas, resources, outcomes, and problems related to use of the innovation.
LEVEL 0 NON-USE: State in which the user has little or no knowledge of the innovation, no involvement with the innovation, and is doing nothing toward becoming involved.	Knows nothing about this or similar innovations or has only very limited general knowledge of efforts to develop innovations in the area.	Takes little or no action to solicit information beyond reviewing descriptive information about this or similar innovations when it happens to come to personal attention.	Is not communicating with others about the innovation beyond possibly acknowledging that the innovation exists.
DECISION POINT A	*Takes action to learn more detailed information about the innovation.*		
LEVEL I ORIENTATION: State in which the user has acquired or is acquiring information about the innovation and/or has explored or is exploring its value orientation and its demands upon user and user system.	Knows general information about the innovation such as origin, characteristics, and implementation requirements.	Seeks descriptive material about the innovation. Seeks opinions and knowledge of others through discussions, visits, or workshops.	Discusses the innovation in general terms and/or exchanges descriptive information, materials, or ideas about the innovation and possible implications of its use.
DECISION POINT B	*Makes a decision to use the innovation by establishing a time to begin.*		
LEVEL II PREPARATION: State in which the user is preparing for first use of the innovation.	Knows logistical requirements, necessary resources and timing for initial use of the innovation, and details of initial experiences for clients.	Seeks information and resources specifically related to preparation for use of the innovation in own setting.	Discusses resources needed for initial use of the innovation. Joins others in pre-use training, and in planning for resources, logistics, schedules, etc., in preparation for first use.
DECISION POINT C	*Begins first use of the innovation.*		
LEVEL III MECHANICAL USE: State in which the user focuses most effort on the short-term, day-to-day use of the innovation with little time for reflection. Changes in use are made more to meet user needs than client needs. The user is primarily engaged in a stepwise attempt to master the tasks required to use the innovation, often resulting in disjointed and superficial use.	Knows on a day-to-day basis the requirements for using the innovation. Is more knowledgeable on short-term activities and effects than long-range activities and effects of use of the innovation.	Solicits management information about such things as logistics, scheduling techniques, and ideas for reducing amount of time and work required of user.	Discusses management and logistical issues related to use of the innovation. Resources and materials are shared for purposes of reducing management, flow and logistical problems related to use of the innovation.
DECISION POINT D-1	*A routine pattern of use is established.*		
LEVEL IV A ROUTINE: Use of the innovation is stabilized. Few if any changes are being made in ongoing use. Little preparation or thought is being given to improving innovation use or its consequences.	Knows both short- and long-term requirements for use and how to use the innovation with minimum effort or stress.	Makes no special efforts to seek information as a part of ongoing use of the innovation.	Describes current use of the innovation with little or no reference to ways of changing use.
DECISION POINT D-2	*Changes use of the innovation based on formal or informal evaluation in order to increase client outcomes.*		
LEVEL IV B REFINEMENT: State in which the user varies the use of the innovation to increase the impact on clients within immediate sphere of influence. Variations are based on knowledge of both short- and long-term consequences for clients.	Knows cognitive and affective effects of the innovation on clients and ways for increasing impact on clients.	Solicits information and materials that focus specifically on changing use of the innovation to affect client outcomes.	Discusses own methods of modifying use of the innovation to change client outcomes.
DECISION POINT E	*Initiates changes in use of innovation based on input of and in coordination with what colleagues are doing.*		
LEVEL V INTEGRATION: State in which the user is combining own efforts to use the innovation with related activities of colleagues to achieve a collective impact on clients within their common sphere of influence.	Knows how to coordinate own use of the innovation with colleagues to provide a collective impact on clients.	Solicits information and opinions for the purpose of collaborating with others in use of the innovation.	Discusses efforts to increase client impact through collaboration with others on personal use of the innovation.
DECISION POINT F	*Begins exploring alternatives to or major modifications of the innovation presently in use.*		
LEVEL VI RENEWAL: State in which the user re-evaluates the quality of use of the innovation, seeks major modifications of or alternatives to present innovation to achieve increased impact on clients, examines new developments in the field, and explores new goals for self and the system.	Knows of alternatives that could be used to change or replace the present innovation that would improve the quality of outcomes of its use.	Seeks information and materials about other innovations as alternatives to the present innovation or for making major adaptations in the innovation.	Focuses discussions on identification of major alternatives or replacements for the current innovation.

Procedures for Adopting Educational Innovations Project, Research and Development Center for Teacher Education, University of Texas at Austin, 1975, N.I.E. Contract No. NIE-C-74-0087.

BEST COPY AVAILABLE

FIGURE 1 — LoU CHART

CATEGORIES

ASSESSING	PLANNING	STATUS REPORTING	PERFORMING
Examines the potential or actual use of the innovation or some aspect of it. This can be a mental assessment or can involve actual collection and analysis of data.	Designs and outlines short- and/or long-range steps to be taken during process of innovation adoption, i.e., aligns resources, schedules activities, meets with others to organize and/or coordinate use of the innovation.	Describes personal stand at the present time in relation to use of the innovation.	Carries out the actions and activities entailed in operationalizing the innovation.
Takes no action to analyze the innovation, its characteristics, possible use, or consequences of use.	Schedules no time and specifies no steps for the study or use of the innovation.	Reports little or no personal involvement with the innovation.	Takes no discernible action toward learning about or using the innovation. The innovation and/or its accouterments are not present or in use.
Analyzes and compares materials, content, requirements for use, evaluation reports, potential outcomes, strengths and weaknesses for purpose of making a decision about use of the innovation.	Plans to gather necessary information and resources as needed to make a decision for or against use of the innovation.	Reports presently orienting self to what the innovation is and is not.	Explores the innovation and requirements for its use by talking to others about it, reviewing descriptive information and sample materials, attending orientation sessions, and observing others using it.
Analyzes detailed requirements and available resources for initial use of the innovation.	Identifies steps and procedures entailed in obtaining resources and organizing activities and events for initial use of the innovation.	Reports preparing self for initial use of the innovation.	Studies reference materials in depth, organizes resources and logistics, schedules and receives skill training in preparation for initial use.
Examines own use of the innovation with respect to problems of logistics, management, time, schedules, resources, and general reactions of clients.	Plans for organizing and managing resources, activities, and events related primarily to immediate ongoing use of the innovation. Planned-for changes address management or logistical issues with a short-term perspective.	Reports that logistics, time, management, resource organization, etc., are the focus of most personal efforts to use the innovation.	Manages innovation with varying degrees of efficiency. Often lacks anticipation of immediate consequences. The flow of actions in the user and clients is often disjointed, uneven and uncertain. When changes are made, they are primarily in response to logistical and organizational problems.
Limits evaluation activities to those administratively required, with little attention paid to findings for the purpose of changing use.	Plans intermediate and long-range actions with little projected variation in how the innovation will be used. Planning focuses on routine use of resources, personnel, etc.	Reports that personal use of the innovation is going along satisfactorily with few if any problems.	Uses the innovation smoothly with minimal management problems; over time, there is little variation in pattern of use.
Assesses use of the innovation for the purpose of changing current practices to improve client outcomes.	Develops intermediate and long-range plans that anticipate possible and needed steps, resources, and events designed to enhance client outcomes.	Reports varying use of the innovation in order to change client outcomes.	Explores and experiments with alternative combinations of the innovation with existing practices to maximize client involvement and to optimize client outcomes.
Appraises collaborative use of the innovation in terms of client outcomes and strengths and weaknesses of the integrated effort.	Plans specific actions to coordinate own use of the innovation with others to achieve increased impact on clients.	Reports spending time and energy collaborating with others about integrating own use of the innovation.	Collaborates with others in use of the innovation as a means for expanding the innovation's impact on clients. Changes in use are made in coordination with others.
Analyzes advantages and disadvantages of major modifications or alternatives to the present innovation.	Plans activities that involve pursuit of alternatives to enhance or replace the innovation.	Reports considering major modifications of or alternatives to present use of the innovation.	Explores other innovations that could be used in combination with or in place of the present innovation in an attempt to develop more effective means of achieving client outcomes.

APPENDIX K

Inservice Training Component
Developmental Writing Program

DEVELOPMENTAL WRITING

Inservice Points

60 maximum, over a multi-year period

General Objectives

To provide instruction to elementary classroom teachers in techniques of teaching writing as a developmental process.

Specific Objectives

1. The participant will teach writing as a process involving the following steps:

   a) pre-writing,
   b) drafting,
   c) three levels of editing, i.e., self, peer, teacher,
   d) revising, and
   e) publishing/sharing.

2. The participant will use writing instruction techniques appropriate for the elementary level, e.g., sentence expansion and paragraph development.

3. The participant will evaluate student's writing using a variety of techniques.

4. The participant will systematically implement the developmental writing program (DWP) in his/her setting according to program definition and guidelines.

5. The participant will develop:

   a) a positive attitude towards teaching writing in general and towards his/her own writing in particular, and
   b) a commitment to further acquisition of personal writing skills.

6. Learners for whom the participant is responsible will demonstrate improvements in:

   a) general writing production,
   b) language mechanics and expression,
   c) critical reading skills, and
   d) attitudes toward writing.

Description of Activities

During the course of the initial training year, participants will receive 18 hours of group instruction covering topics listed in objectives 1-3. All

participants will be provided with materials necessary to implement the program, and trained in their use. Supervisory staff who are directly responsible for this component will provide ongoing support and assistance in program implementation to all participants through systematically scheduled site visitations, observations, and conferences.

During subsequent training years, participants will receive 12 hours of group instruction delivered throughout the course of each academic year.

Materials specific to the needs of individual participants will be provided. Supervisory staff will continue to provide ongoing support and assistance on a systematic basis.

Evaluation

Because this is an originally-designed local writing program, it is assumed that participants have little or no knowledge of program specifics prior to their involvement in the training program.

Evaluative activities will focus on changes in actual teaching behaviors, teacher attitude, and student outcomes.

Program Developer(s)

Jackie Matthews, Elementary Supervisor
George Mitchell, Elementary Supervisor

ATTACHMENT 1

Developmental Writing Program Scoring Manual

SCORING MANUAL

DEVELOPMENTAL WRITING PROGRAM

HILLSBOROUGH COUNTY SCHOOL BOARD
-1984-

Developed By:
Kathryn Kaney
PDE Associates, Inc.

TABLE OF CONTENTS

INTRODUCTION

ISSUES IN HANDSCORING WRITING SAMPLES

SCORING PROCEDURES AND REGULATIONS

EXPECTATIONS FOR THE COMPOSITION

FOCUSED HOLISTIC SCORE SCALE

ANCHOR PAPERS AND RATIONALE

INTRODUCTION

The Hillsborough County Developmental Writing Program (DWP) was piloted in eleven district elementary schools during the 1982-83 school year and was expanded to twenty-three additional schools for the 1983-84 school year.

The program's purposes are:

to identify an effective model for systematically developing writing instruction for selected students in grades 4-6;

to fieldtest Teaching Writing: A Developmental Systematic Approach, authored by Evelyn Rothstein, as the basis for teacher training and subsequent classroom instruction;

to integrate presently adopted materials and identify additional resources needed to maintain an ongoing developmental writing program; and

to determine the advisability of expanding the instructional model to include all elementary schools.

Most of the twenty teachers who participated in the program in 1982-83 have continued in the program during 1983-84. In addition, ninety-two new teachers have joined the program this year, and the program has been expanded to twenty-three new schools. Targeted grade levels remain fourth, fifth, and sixth, as they were during the 1982-83 school year.

This scoring activity is part of a larger evaluation effort to determine how effective the Developmental Writing Program is. Results of the scoring activity will be used to provide answers to perhaps the most important of all the research questions:

Do students who receive instruction in the Developmental Writing Program write better than students who participate in other structured and unstructured writing programs?

To provide as much information as possible in answering this question, the Department of Testing and Evaluation has designed a comprehensive evaluation which includes examining change among multiple experimental groups as well as a control group.

Your participation in this scoring activity is critically important to the integrity of the evaluation results, and we appreciate your enthusiastic involvement. You will receive extensive training for handscoring writing samples, and we hope that the training will continue to be valuable to you long after our formal scoring activities are over.

ISSUES IN HANDSCORING WRITING SAMPLES

Most hand-scoring procedures currently in use throughout the country can be classified under two rubrics: holistic and analytical. Holistic scoring requires that the Reader evaluate the overall quality of a piece by permitting all characteristics of the piece to interact with each other in influencing his/ her judgment. Analytical scoring requires that the Reader isolate each particular characteristic under examination and judge each without being influenced by any of the others.

Decisions about which of these two approaches should be used in a scoring effort must be made in light of the type of information that one needs or desires about a program. The information that is needed about this program--whether youngsters who have been in the program write better than youngsters who have not--requires a judgment of overall quality. For that reason, a scoring procedure of the holistic type has been chosen for use.

Pure holistic scoring has certain measurement properties which cause it to be inappropriate for this evaluation, however. Briefly, because pure holistic scoring is normative in nature, it does not provide the kind of information about performance levels that this evaluation effort requires. To acquire the necessary information while maintaining the assessment of overall quality, a variation of holistic scoring, called focused holistic scoring, will be used. Focused holistic scoring is the most widely adopted of all hand-scoring models currently used in large-scale scoring programs. Its popularity can be attributed to its easy adaptability to program purposes and to the instructional value inherent in the results that it yields.

When hand-scoring activities are used to make significant decisions about examinees or to provide data in a formal research study, certain controls must be imposed. Many of the procedural controls imposed during preparation and during actual scoring are designed with two objectives in mind: that the scores be valid and that they be reliable. Score validity reflects how accurate the scores are. Score reliability reflects how consistent the Readers are.

A score is considered valid if it accurately reflects the examinee's actual performance, or the examinee's "true score." In a performance testing program such as a writing assessment, "true score" is determined by expert judgment. Score validity is monitored throughout scoring by periodically comparing Reader scores with the experts' "true scores," assigned during pre-scoring. Before the scoring session convenes, a group of experts is assembled to conduct pre-scoring activities. During pre-scoring, these experts, who represent the school system's present standards, undergo intensive training. After they have been trained, they score a large number of papers. This scoring is followed by extensive discussion and re-scoring, until the expert group is in agreement about the score assigned to each paper. Some of these pre-scored papers become Anchor Papers, so called because they anchor or stabalize the scale. The Anchor Papers are included in the Scoring Manual and are used to train Readers. Other pre-scored papers are used to monitor Reader accuracy at periodic intervals throughout the scoring session. These are called Check Papers because they are used to check whether Readers are continuing to score in concert with the designated criteria or are beginning to drift away from these criteria.

When it is time to conduct an accuracy check, each Reader is given copies of the same set of papers. When Readers have scored these papers, each Reader's scores are compared with the scores assigned by the experts. Each Reader is given an Accuracy Rating, and these ratings are returned to the Reader, along with each paper's expert or "true score" and an explanation for the "true score." Readers whose Accuracy Rating falls within an acceptable range are asked simply to review the "true scores" and the rationale statements. Readers whose Accuracy Ratings are not within the range of tolerance are required to undergo further training before continuing in formal scoring activities.

A test is considered to be reliable if it consistently yields the same indication of examinee achievement, knowledge, or performance. Several different procedures often are used to evaluate how reliable a test is. Score reliability is evaluated by examining a property closely related to internal consistency and it is attained when two or more observations of performance agree with or are the same as each other. Costs associated with hand-scoring often prohibit our

judging examinee performance on the basis of more than one writing sample, however. In the case of hand-scoring, therefore, we define observations not as items, but as ratings, each Reader-assigned score constituting an observation.

To ensure that ratings are reliable, another type of monitoring is conducted throughout scoring activities. This monitoring is not periodic, as in the case of the accuracy checks. Rather, it is on-going. Each paper is scored by two or more different Readers. To guarantee that each score is independent, the first score is obscured from the view of subsequent Readers. Once a set of papers has been scored by all of its Readers, the scores on each paper are compared. The more often all of the scores are the same, the more reliable the scores are judged to be. Just as some deviation can be tolerated in the Accuracy Rating, some deviation also can be tolerated in the Agreement Rating. If a set of Readers begins to stray too far from perfect agreement, however, their most recent Accuracy Ratings are checked to determine whether one or more of the Readers is inaccurately applying the score scale. The inaccurate Reader or Readers is then removed from formal scoring and retrained until they have regained sufficient Accuracy to be reinstated.

Large-scale scoring activities such as this one require enormous planning and organization efforts both before and during the formal scoring session, not only to ensure that scores are accurate and reliable, but to be sure that the paper flow is smooth, that all scores are independent, and that scoring is reasonably distributed across all Readers. From time to time during the scoring session, Readers are asked to wait quietly while the scoring staff completes calculations necessary for monitoring accuracy and consistency. Patient and cooperative Readers always are appreciated.

SCORING PROCEDURES AND REGULATIONS

Because each of you has a background that includes some study of and/or experience in writing, all of you share some knowledge and understanding in common. The training and experience that you bring to this scoring activity will, for the most part, serve you well. They will enable you to read and score compositions with an informed eye.

Just as each of you brings special knowledge and experience to this activity, however, each also brings special biases and prejudices about the relative importance and appropriateness of certain characteristics of written composition. These prejudces and peeves can vary significantly from one Reader to another.

The first objective of training is to capitalize on the common knowledge and understanding among Readers, and to remove, at least temporarily, the biases and special peeves that they might hold. The training material is designed to give clear definition to each characteristic that you will be asked to evaluate. In every instance where possible, the Scoring Guide reduces judgement and inference to the lowest possible level. Even so, judgement and inference are necessary. It will be essential, therefore, that you accept the definitions set forth in this Guide, notwithstanding your own feeling that you might assign a different score if you could use your own standards to rate the composition. The tension between your own standards and those imposed in this Guide will frustrate you from time to time. It might be of some comfort to you to know that everyone who ever was involved in a scoring activity like this one has experienced all of the frustrations that you are about to experience and that all of them have survived.

Each composition will be scored by two Readers. Our expectation is that the score assigned by each Reader will be the same in almost all cases. Because each of you will score compositions in common with two other Readers, it is important that all of you reach a common understanding of the definitions and of how to apply them. Your consistent agreement with other Readers and your accurate adherence to the scoring criteria will be monitored throughout the

scoring session. If it becomes apparent that your scoring has become inconsistent or inaccurate, you will be required to undergo retraining exercises.

The Trainer will answer your procedural questions and will clarify scoring criteria, if necessary. The Trainer will not, however, advise you about the score that a composition should receive.

Several rules must be observed during the scoring session:

1. Each Reader's score must be the result of an independent evaluation of the composition. You will not be allowed to confer with another Reader regarding a composition. If you are tempted to share something particularly delightful with a fellow Reader, suppress the urge and read on.

2. It is important that a quiet atmosphere be maintained at all times. If you need clarification or assistance, raise your hand and the Trainer will come to you. In conferring with the Trainer, speak as quietly as possible so that other Readers will not be distracted by your conversation.

3. Scoring efficiency is important. You will need to concentrate on each composition so that one reading will allow you to assign a score without re-reading or re-scanning the composition. Reasonable speed is important, so concentrate and establish an efficient pace.

4. If you need to take a break to stretch your legs, you may quietly walk around in the designated area of the scoring room, or you may leave the scoring room for a few minutes. Two 15-minute breaks will be provided each day, one in the morning and one in the afternoon. In addition, you will be given a one-hour lunch break. You will be expected to accomplish 6½ hours of actual scoring each day.

5. Prolonged near-point focus can cause eye fatigue. If you begin to experience tiredness or burning in your eyes, take a few minutes to look up and focus your eyes on a distant point. This will help to prevent headaches and blurred vision.

6. Under no circumstance should you remove scoring materials from the scoring room. When you take a break, you should leave all of your materials on the table in front of you.

Your after-hours activities are not governed by this project. Nonetheless. you are encouraged to rest your eyes and get a good night's sleep each night.

EXPECTATIONS FOR THE COMPOSITION

Both of the prompts used for this program are expected to elicit descriptive discourse from examinees. Although the central composing effect is expected to be descriptive, students may choose either an explanatory or a narrative explication strategy. Even though explicating strategies may vary among writers, the same criteria will be used to score all papers.

The score scale refers to performance in four categories or domains: composing, sentence formation, usage, and mechanics. A general discussion of each domain is provided below.

COMPOSING: The Composing Domain comprises all of the characteristics of a composition that influence the effective delivery of the writer's message to achieve the writer's purpose. These characteristics include specifying and focusing on the subject-matter, providing information that explicates the subject-matter, and delivering the discourse through organized, logical, and coherent discourse.

SENTENCE FORMATION: The Sentence Formation domain comprises all of the characteristics of composition that influence the effective expression of assertions, propositions, questions, commands, and exclamations such that T-units and their constituents are internally related and are either separated from or coordinated with each other. These characteristics include variety in sentence length and structure, and the absence of formation errors such as run-ons, fragments, misplaced modifiers, faulty pronoun references, dangling modifiers and nonparalled elements.

USAGE: The Usage Domain comprises all of the characteristics of language selection and use that cause a composition to be acceptable and effective for a standard, informal, discourse intended for an audience committed to national, reputable, and present standards. These characteristics include the selection of language appropriate for the purpose of the piece, the presence of

agreement in number between subject noun and verb elements, the agreement in number, gender and case between nouns and pronouns, and the use of correct and appropriate inflected forms.

MECHANICS: The Mechanics Domain is the only one of the four domains which is restricted exclusively to written language. This domain comprises the system of symbols and cuing devices that allow us to represent language in written text. The symbol system includes punctuation, capitalization, and the graphemic symbols that we commonly refer to as spelling. The translatability, or legibility of these symbols also is classified as a skill in the Mechanics Domain. Cuing devices such as paragraph indentation and text formatting also are members of the Mechanics Domain.

Many of our expectations about effective composition are the same, regardless of the composition's type or the author's purpose. We expect all compositions to be focused, developed, organized, and logical. We expect them to be free of sentence formation errors, to communicate through use of standard American English, and to be free of mechanical errors. All of these traits being present and equal, however, we often find that real differences in effectiveness rest in he writer's explication strategy and language selection. In responding to Prompt A, writers are expected to select a delivery strategy that will either describe the pet that they would like to have or express reasons for wanting to have the pet that they identify. The language that the witer selects should assist him/her in relating his/her wishes to the reader.

In responding to Prompt B, writers are expected to select a delivery strategy that will describe a fear, express reasons for a fear, or narrate an event that caused them fear. Once again, the writer is expected to use language that will assist him/her in relating his/her fear to the reader.

FOCUED HOLISTIC
SCORE SCALE

A five-point focused holistic scale will be used in this scoring activity. Only four of the scale points, 1 through 4, will be used to judge the quality of performance, however. The zero value will be reserved for certain anomolous responses and it cannot be used unless permission has been given by the Trainer. Each scale value is defined by an anchor statement and by several actual writing samples which are judged to illustrate the specified level of performance. The scale points and their respective Anchor statements are given below. Each is restated in the following chapter, where Anchor Papers and rationale statements are provided to further explicate each scale point.

4 = The composition is clearly focused and the subject-matter is well-developed. The writer's delivery is well-organized, logical, and coherent. Sentences are correctly formed, and the writer has varied sentence structure and length. The writer's language is effective and is appropriate for standard, informal written discourse. The writer shows evidence of a consistent command of written language mechanics.

3 = The composition is reasonably focused and the subject-matter is adequately developed. The writer's delivery is sufficiently organized, logical, and coherent. Most sentences are correctly formed, and the writer's general command of sentence formation is evident. The writer has a reasonable variety of sentence structure and length. Overall, the writer's language is reasonably effective and is acceptable for standard, informal written discourse, though minor deviations may exist. The writer shows evidence of an overall understanding of written language mecanics; deviations are minor.

2 = The composition's focus is weak. Although the subject-matter is discernible, development is minimal. Organization, logic, and coherence are minimal, but present. Some understanding of sentence formation is evident, but deviations may predominate. If the writer has a variety of sentence structure and length, it may be limited. Ineffective and inappropriate language use may cause the Reader to question the writer's understanding of standard, informal written discourse. Some knowledge of written language mechanics is evident, but may be weak.

1 = The composition's focus is very weak, if present. The subject-matter may be marginally discernible, but is undeveloped. The composition is marred by weaknesses in organization, logic, and coherent delivery. Sentence formation errors are obtrusive, and the composition has little if any variety in sentence structure and length. Deviations from language expected in standard, informal written discourse are major, and command is not evident. The writer shows little if any evidence of an understanding of written language mechanics.

0 = The composition cannot be scored. YOU MUST RECEIVE THE TRAINER'S PERMISSION TO ASSIGN A "0".